



**EES 2018 GREECE**

THESSALONIKI, 1–5 OCTOBER

**13<sup>th</sup> European Evaluation Society  
Biennial Conference**

# **Evaluation for More Resilient Societies**



## **Abstract Book**

1–5 October 2018

Ioannis Vellidis C.C., TIF HELEXPO SA

Thessaloniki, Greece

[www.ees2018.eu](http://www.ees2018.eu)

# Content

## Keynote Lectures 25

K 1 - Resilient Societies and Resilient Evaluators in Turbulent Times .....	26
K 2 - Post-normal Evaluation? .....	27
K 3 - Caroline Heider's Keynote Lecture .....	28
K 4 - Learning About Evaluation from a Conference? .....	29

## Flagship Symposia 30

Better Evaluation for Better Policies in Greece Beyond the Crisis .....	31
FS 01 - Contribution of Ms. Gerovasilis .....	32
FS 02 - Contribution of Ms Louka Katseli .....	32
FS 03 - Contribution of Ms Anastasia Tzavaras Catsambas .....	32
Evaluation Unusual: How is Resilience Challenging the Evaluation Field, and Are We Rising Effectively to the Challenge? .....	33
FS 04 - Contribution of Tim Frankenberger .....	34
FS 05 - Contribution of Jyotsna Puri .....	34
FS 06 - Contribution of Maliha Khan .....	34
FS 07 - Contribution of Claire Hutchings .....	34
Evaluation Criteria for the 21 <sup>st</sup> Century: Panel 2 Revisiting the DAC Criteria .....	35
FS 08 - N/A - Interactive Discussion .....	36
FS 09 - N/A - Interactive Discussion .....	36
FS 10 - N/A - Interactive Discussion .....	36
FS 11 - N/A - Interactive Discussion .....	36
FS 12 – Symposium Contribution .....	38
FS 13 – Symposium Contribution .....	38
FS 14 – Symposium Contribution .....	38
FS 15 - Symposium Contribution .....	38

## Fishbowl Sessions 39

Promoting Gender Responsive and Equity Focused Evaluation among Young and Emerging Evaluators: Experiences in Cross Collaboration between EvalYouth and EvalGender+ .....	40
FB 01 - It is Never Too Early or Too Late to Learn About EFGRE: Co-Chairs Experiences with Capacity Development .....	41
FB 02 - Contribution of Rituu B. Nanda .....	41
FB 03 - A Local Perspective of Evalgender and Evalyouth Collaboration; The Latin American Experience .....	41
EVALSDDGs - Evaluation is an Essential Ingredient for the SDGs – Sharing Experiences .....	43
FB 08 - Contribution of Dorothy Lucks .....	44
FB 09 - Contribution of Stefano D'Errico .....	44
FB 10 - Contribution of Kassem ElSaddik .....	44
FB 11 - Contribution of Ada Ocampo .....	44
Evaluation Confidential – Adventures from the Evaluator's Kitchen .....	45
FB 12 - Introduction to NIRAS' Evaluation Approach, Process and Examples .....	46

FB 13 - Implications of Allocated Evaluation Resources – Juggling Between Technical Time Versus Management Time and Finding the Right Balance Between Time for Familiarisation Versus Time for Evaluation . . . . .	47
FB 14 - Stakeholders Management – Responding to Stakeholders' Requests and Learning Interests while Managing Expectations and Ensuring the Evaluator's Independence . . . . .	48
The Work of the European Commission Regulatory Scrutiny Board . . . . .	49
FB 15 - Contribution of Vesna Valant . . . . .	49
FB 16 - Contribution of Joakim Nilsson . . . . .	49
FB 17 - Contribution of Michael White . . . . .	49
FB 18 - Contribution of Guenther Ebling . . . . .	49

## Special Session

50

SS 04 Panel on Regional VOPEs . . . . .	51
SC 03 – Panel Contribution . . . . .	51
SC 04 – Panel Contribution . . . . .	51
SC 05 - Panel Contribution . . . . .	51
SC 06 - Panel Contribution . . . . .	51
SC 07 - Panel Contribution . . . . .	51
SC 08 - Panel Contribution . . . . .	51

## Paper Sessions

52

O 001 - Approaches for Making Use of Unstructured Big Data in Evaluation . . . . .	53
O 002 - Can the Use of Big Data Overcome the Limits of Large Scale Surveys and Other Types of Stakeholder Consultations? . . . . .	54
O 003 - Collateral Damages by Shotgun Shooting: Why Big Data May Lead to Big Trouble . . . . .	55
O 004 - Is it Possible to use Big Data for UN Women Evaluations? . . . . .	56
O 005 - How Firms Evaluate Their Contribution to the Sustainable Development Goals: Tools, Gaps and Future Challenges . . . . .	57
O 006 - Global Evaluation Agenda 2016–2020, Bite of EvalApple? . . . . .	58
O 007 - Strengthening Institutions and National Evaluation Capacity for Achieving the SDGs: Increased Accountability, Transparency, and Public Policy Evaluation in Morocco . . . . .	59
O 008 - Is Better Regulation any Good for European Citizens? . . . . .	60
O 009 - Evaluation of an EU Funding Programme Which, Due Turbulent Times, Took a Very Different Form . . . . .	61
O 010 - Evaluating Gender Mainstreaming in European Research & Innovation Policy: Interim Evaluation of Gender as a Cross-Cutting Issue in Horizon 2020 . . . . .	62
O 011 - Using Evaluation to Understand the Potential to Scale Up Resilience Innovations . . . . .	63
O 012 - The Role of Monitoring and Evaluation for Global Climate Resilience Programmes: Challenges, Lessons and Ways Forward . . . . .	64
O 013 - Coping with The Tragedy of the All Too Common: Common Themes in Evaluating Climate Change Mitigation . . . . .	65
O 015 - Challenges in Institutional Set-Up After Devolution from Federal to Provincial Government – A Unique Case of PWD . . . . .	66
O 016 - Systematic Integration of Meta-Evaluation and Evaluation Synthesis: Aggregating Evidence Through a Quantitative Evaluation Synthesis Design . . . . .	67
O 017 - The Challenges of Screening and Synthesising Qualitative Research in a Mixed-Methods Systematic Review on the Impact of Agricultural Certification Schemes . . . . .	68
O 018 - Knowledge Accumulation in Systematic Reviews: Advancing Strategies for Causal Generalization . . . . .	69
O 019 - Bridging The Gap Between Innovation Support and Socio-economic Impacts in Green Technology . . . . .	70
O 020 - Evaluation Methodology – Critical Foundations from Finland . . . . .	71
O 021 - Human-Centered Co-Evaluation as a Lever for Learning and Evaluation Capacity Building (ECB) . . . . .	72

O 022 - Overcoming Poverty and School Desertion. The Mixed Methods Evaluation Process of a “Programa Prospera” Pilot Program in Mexico .....	73
O 023 - Impact of Poverty Alleviation Through Unconditional Cash Transfers-Benazir Income Support Program in Pakistan .....	74
O 024 - Focusing on Double Vision: Are Proxy Means Tests Effective to Identify Future School Dropouts and the Poor? .....	75
O 025 - Evaluators Unite: How RTI International has Leveraged Evaluation Expertise in Times of Critical Funding Constraints and Demand for Results .....	76
O 026 - Learning for Adapting: Findings and Implications of a Peer Education Assessment in the Philippines .....	77
O 027 - Can Social Impact Evaluations be Standardized? Results from Introducing a Toolkit to Dementia Challenge Participants .....	78
O 028 - Designing M&E in a Dynamic World: A Case of African Dynamic Environment .....	80
O 029 - Towards a General Theory of Monitoring and Evaluation: The Measurement Of Organisational Capacity .....	81
O 030 - Is Evaluating and Improving Leadership by Using Quality Standards and Indicators a “Hollow Bunny Trauma” .....	82
O 031 - The Impact of Education Programmes, Evidence from an Evaluation of the Integrated School Performance Improvement, Review and Engagement Project, Malawi .....	83
O 032 - School Self-evaluation Capacity Building – Should We Model it? .....	84
O 033 - Reconsidering Interventions as Time-Varying Events: Lessons Learned from an Evaluation of Early Grade Retention .....	85
O 034 - Adoption and Impacts of Improved Beekeeping Technologies in the Miombo Woodland of Tanzania .....	86
O 035 - Evaluator Competencies and University-Based Evaluation Education In 2018: Global Perspectives and Implications for Professionalism .....	87
O 036 - Training Evaluators to be Resilient and Action Oriented: Contrasting Perspectives on Teaching .....	88
O 037 - Professional Networks and Professionalization in the Evaluation of Public Policies in Ibero-America .....	89
O 038 - Crossing Evaluation Boundaries in the Face of Uncertainties: A Taxonomy to Commission, Conduct and Use Better Quality and Resilient Evaluations .....	90
O 039 - Revisiting the Lessons of the Global Decolonisation Movement for More Effective Research and Evaluation .....	91
O 041 - Evaluators and VOPEs With Resilient Cities: the ABC of What to Know and Do from Day 0 .....	92
O 042 - Using Theory of Change Models to Promote Inclusive Climate Action in Cities – the Urban Climate Action Impact Framework (UCAIF) .....	93
O 043 - Tracking Contextualised Adaptation Across Scales: The TAMD Framework .....	94
O 044 - Opportunities and Challenges for Using Performance Measurement Information in Support of Resilience: The Case of Local Governments in Greece .....	95
O 045 - A Practical Approach to Address the Attribution Challenge in Complex Interventions: Evaluating Impact of a Market Systems Development Programme .....	96
O 047 - Contribution Analysis and Translation Theory to Measure Use and Influence of a Evaluation: Rethinking Evaluation Utility Standards .....	97
O 048 - Ethics in Evaluation; A Factor of Resilience? Whose Resilience? .....	98
O 049 - The Other Side of Gender Inequality: Men and Masculinities in Afghanistan .....	99
O 050 - Women Leaders Project, Madrid City Council. An Evaluation Based on the Theory of Change With Gender & Human Rights Approach .....	100
O 051 - Institutionalisation of Policy Evaluation as Enabler for Sound Public Governance: Towards an OECD Perspective .....	101
O 052 - The Influence of Evaluations in Polycentric Settings: a Literature Review .....	102
O 053 - Can Evaluations Really Contribute to Evidence-Based Policy Making at Government Level? – The Case of the French Government Modernisation Evaluations .....	103
O 054 - Why Does Europe Leave SDG Evaluation Behind? .....	104
O 055 - SDGs National Reports, the Gap Between the Reality and What Governments Show. A Review of Latin-American Reports About SDG 5 .....	105
O 056 - Sustainability Reporting, SDGs, UN-Global Compact and GRI: How is Private Sector Communicating its Impact? .....	106
O 057 - Using a System Approach to Rationalise Policy Evaluation in the Field of Science Communication and STEM (Science, Technology, Engineering, Mathematics) .....	107
O 058 - Assessing Impact of Taiwan’s Publicly-Funded STI Program on Participant Firms: the Empirical Evidence and Policy Implications .....	108
O 059 - When the Measurement Explains the Effect: The Effect of Baseline Survey Participation on Intercultural Training Effectiveness Outcomes .....	110
O 060 - Evaluation of Non-Formal Education for Mental Health Professionals: The Case of the Appreciative Systemic Inquiry Experiential Educational Program of OKANA .....	111

O 061 - Is Training Evaluation a Threatened Discipline? The Dilemma of Theory Vs. Practical Reality .....	112
O 062 - Resilience Through Adaptation: The Role of Evaluation in Driving Reflection and Change in Adaptive, Democracy Assistance Programmes .....	113
O 063 - Knowledge-Focused Evaluation in Changing Contexts – What Evaluation Synthesis can Bring to the Table .....	114
O 064 - Evaluation as Therapy: An Evaluation of Corporate – NGO Partnership between C&A, C&A Foundation and Save the Children .....	115
O 066 - How to Strengthen Our Coordination for Resilience of the Poorest: A Theory of Change & Learning from the Ethiopian Experience .....	116
O 067 - Usage of Evaluation Tools to Identify Causes and Trends of Quasi Bellic Conflicts in Latin America (Venezuela, Ecuador and Bolivia) .....	118
O 068 - Evaluating Emotion: The Use of Counseling Skills in Responsive Qualitative Evaluation .....	119
O 069 - Participatory Processes: When the Evaluated Become the Evaluators .....	120
O 070 - Once Upon A (Different) Time: The Value of Story in a Tempestuous World .....	121
O 071 - Can Evaluation Save Results-Based Management? .....	122
O 072 - Time-Driven Activity-Based Costing and Monitoring and Evaluation Systems in the Public Sector. The Case of the General Chemical State Laboratory .....	123
O 073 - Evaluation Methodology for Post-Legislative Scrutiny? .....	124
O 074 - "Traditional" Approaches in the Age of RCTs and Big Data .....	125
O 075 - Rethinking Evaluation in US Higher Education: Meeting Evaluation Needs in Rapidly Shifting Climates .....	126
O 076 - A Comprehensive Framework to Design and Conduct Evaluation .....	127
O 077 - Is Evaluator-Focused Meta-Evaluation Occurring or Just Theorized? .....	128
O 079 - Evidence Summaries For Policy Making – Using A Scoping Review To Inform Implementation Of The SDG14 Fisheries Targets	129
O 080 - Embedding Ex-Ante Evaluation of New Activities in an EU Agency – Some Lessons Learned .....	130
O 081 - This Paper will Explore how the Programming of EU Agencies' Work and Programmes can be Improved .....	132
O 082 - Gendering Evaluation – Integration of Gender in Evaluation Standards and Alternative Approaches .....	133
O 083 - Putting the "Theory" Back into Theory-Based Evaluation: Making the Case for Theory-Knitting in Evaluation .....	134
O 084 - Like Cat and Dog? Mixing Counterfactual and Contribution Analyses in Evaluating Complex Policies .....	135
O 085 - Innovations in Theory-Based Evaluation: Using Nested Actor-Based Theories of Change to Evaluation a Complex International Development Intervention .....	136
O 086 - Migration and Educational Outcomes: Assessing the Impacts of Refugee Status on Learning in Uganda.....	137
O 087 - IKEA Initiatives to address the Refugee and IDP Crisis TWG on Private Sector Evaluation .....	138
O 088 - Challenges to Evaluating Policies and Initiatives Addressing the Migration/Refugee Crisis .....	139
O 089 - Evaluating Social Cohesion to Understand Resilience Among Internally Displaced People Living in Somalia's Urban Areas	140
O 091 - Evaluating Regulatory Impact Assessment in Switzerland .....	141
O 092 - Extending Evaluation Towards Regulatory Impact Assessment to Reduce Regulatory Burdens: A Neo-institutional Study of Context, International Impulses, and Finnish Experiences .....	142
O 093 - Catalyzing Digital Employment Opportunities and Skills Training for African Youth: Assessing Impact Using a Case Study Approach .....	143
O 094 - Evaluating Interventions Targeting Young NEETs in Romania: Data Collection Challenges and Results .....	144
O 096 - Evaluators, Empathy and Policymaking – Can a Better Understanding of the Policymaking Process Enhance the Use of Evidence? .....	145
O 097 - User-Driven Evaluation and Innovation in Recovery Oriented Services in the Welfare and Health Sector .....	146
O 098 - Not About Us Without Us – How Far Do We Go? .....	147
O 099 - Evaluations in Emergency, Recovery and Resilience Context: Lessons from Myanmar, The Philippines, and South Sudan	148
O 100 - What Are We Evaluating For – Better Delivered Programmes Or Increasing Research Bank? .....	149
O 102 - Integrating a Gender Perspective in Country Development Strategies and Programs: Lessons from the World Bank Group's Country Program Evaluation in Rwanda .....	150
O 103 - When Evaluation Meets the Need to Build up Resilient Communities: A Case Study on Women in Fisheries in the EU .....	151
O 104 - Impact Design for Women Empowerment and its Results .....	152
O 105 - Evaluation Governance in a Complex System: A Case Study of CGIAR Evaluations .....	153
O 107 - Using Theory of Change to Structure the Evaluation of the "Dynamics" of the Flemish "Focus on Talent" Policy .....	154

O 108 - Plagiarism Undermines Achievement of Quality Education in African Universities .....	155
O 109 - Towards a General Theory of M&E: An Approach when Evaluating Potentially Hostile and Dishonest Clients .....	156
O 110 - A study on the Evaluation of Ethical Environment for Science and technology .....	157
O 112 - Use of Technology to Improve Planning, Budgeting and Reporting: Lessons from Uganda's Program-Budgeting System (PBS) .....	158
O 113 - Evaluation of Gender Budgeting Enforcement Measures: The Ugandan Experience .....	159
O 114 - Evaluation of Blending as an Aid Modality .....	160
O 115 - Evaluation in the Governance of and Through Social Investments .....	161
O 116 - Handling Complexity to Assess Agricultural Research Impact in Developing Countries: Taking Stock of Lessons from the Impress Ex-Post Method .....	162
O 117 - Evaluating Research Uptake: Methodological Lessons from the Evaluation of the UK Department for International Development's Humanitarian Innovation and Evidence Programme .....	163
O 118 - Developing and Using a Rubric Model for the Evaluation of Prerequisites for Clinical Research in Sweden .....	164
O 119 - How Might a Shift Towards 'Systemic' Evaluation Better Respond to Challenges of Complexity and Uncertainty in Turbulent Times? .....	165
O 120 - Phenomenal Evaluation Culture? Or Responding to Phenomenon-based experimental governance through more systemic and responsive evaluation .....	167
O 121 - A Principles-Based Approach to Evaluating Results at Scale and Scaling Strategies .....	168
O 122 - Combining Relevant Social Theories to Structure Effective Evaluation Practice: The Case of a School to University Transition Program in STEM .....	169
O 123 - Assumptions Based Evaluation .....	170
O 125 - Towards a General Theory of Monitoring and Evaluation: A Standard Theory of Change .....	171
O 126 - Evaluation Systems and Procedures – Is Formalism the Flipside of Institutionalised Evaluation? .....	172
O 127 - Evolution of Croatia's National Evaluation System through Europeanisation .....	173
O 128 - The Role of Evaluation Partnership in Strengthening Mutual Accountability and Development Effectiveness .....	174
O 129 - Drawing Lessons from the Afred Database and the Implications for the Supply and Demand of Evaluations in Africa .....	175
O 130 - Status of Development Evaluation in Ethiopia: A Desk Review .....	176
O 131 - How is it Possible to Design and Conduct Developmental Development Evaluation? .....	177
O 132 - Supporting Responsiveness and Reciprocity: Changing Our Lens from Needs-Based to Rights-Based Evaluation .....	178
O 133 - Common Challenges on Complexity, Intersectionality and Inclusive Evaluation Practices (or Questioning the Myth of the Impartial Evaluation) .....	179
O 134 - Bridging the Divide: The Relationship Between Gender-based Analysis "Plus" (GBA+) Use and its Quality .....	180
O 135 - The Impostor Phenomenon In Evaluation Practice: Scholarly Analysis And Practical Solutions .....	181
O 136 - Integrating the Positive Thinking Approach in Development Evaluation to Deal with Complexity and to Make Local Actors more Resilient .....	182
O 138 - Understanding Usage Complexity in Evaluating Stock Visibility Solution – an mHealth Intervention in South Africa .....	183
O 139 - Expediting Continuous Evaluation Models as Suitable Early Warning Systems for Measuring Impact of Health Programs .....	184
O 140 - The Evaluability of Small-scale Community-based Approaches to Obesity Prevention: Pragmatics, Purpose, Power and Politics .....	185
O 141 - Data Use for Decision and Policy Making Health Facilities in Kenya .....	186
O 142 - Sound Evidence for Timely Decisions in Volatile Contexts Through Impact Evaluation – Experiences of a Large Implementation Organization .....	187
O 143 - Managing for Evaluability and Maximizing Use of Evaluation for More Adaptive Programming: Emergent Challenges and Opportunities from Practice .....	188
O 144 - An Inverted, Project-based Evaluation Classroom for Evaluation Competence Building .....	189
O 145 - Reality Bites: Navigating Quality and Use in a Large-Scale Realist Evaluation .....	190
O 146 - Experience from the Sri Lanka National Evaluation Capacity Building Programme .....	191
O 148 - Exit with Learning. A Tale of Three Perspectives: Evaluation, Learning and Institutional Development .....	192
O 149 - Evaluating Community Ownership: The Advantages of the Feminist Standpoint Approach .....	193
O 150 - The Humanitarian-Development Nexus- What Do Evaluations Have to Say? .....	194
O 151 - Building Bridges Between International Humanitarian and Development Responses to Forced Migration .....	195

O 152 - Impact Evaluation of Interventions for Climate and Livelihood Resilience: Lessons from a Livestock Insurance Scheme in Ethiopia .....	196
O 153 - Large-Scale Global Surveys: Approaches for Data Collection and Validation for a Family Planning Policy Study .....	197
O 154 - Comparison of Livelihood Resources Among Beneficiaries of Critical Ecosystem Management Project in Ajasse-Ipo and Koton-Karfe Watersheds in Nigeria .....	198
O 155 - Re-Examining Evaluation Design to embrace the Sustainable Development Goals: Lessons and Best-fit Practices from Multilateral Organizations .....	199
O 156 - Learning Partnerships: Successfully Generating Purpose-Driven Monitoring and Evaluation (and How to Strike the Balance Between Flexibility, Collaboration and Independence) .....	200
O 157 - From Co-Creation Towards Co-Evaluation: Innovation in Evaluating the Effectiveness of Public Policy Development and Implementation in OGP Action Plans .....	201
O 158 - Simple, Participatory Assessment of Real Change (SPARC): Lessons in Adapting Outcome Harvesting for an HIV Prevention Advocacy Network in Africa .....	202
O 160 - Rethinking Impact Measurement of University Development Cooperation – Strategies for Modular Evaluation Designs .....	203
O 161 - Progressive Evaluation: Challenging Results Driven Culture .....	204
O 162 - Questioning the Relevance of Mainstream Evaluation Approaches in a Complex Environment: An African Story .....	205
O 163 - How Gatekeepers Sometimes Enhance the Effects of Evaluation Machineries in Research .....	206
O 164 - Empirically Exploring if, where, and how Undergraduate Students are Exposed to Evaluation: Implications for Creating Global Pipelines into the Profession .....	207
O 165 - Privilege Cognizant Identity and Evaluation .....	208
O 166 - Closing Civic Space: Formative Research to Strengthen Grant-making Practices and Grantee Resilience .....	209
O 167 - System-Wide Accountability for Gender Equality in the United Nations System .....	210
O 168 - Evaluation of the Inspection and Surveillance Model for Access, Use and Quality in the Provision of Health Services in Colombia .....	211
O 169 - Application of the Constructivist Evaluation to Behavioral Health Issues: Facilitating the Placing of Value by Others .....	212
O 170 - A Novel Methodology to Examine the Know-Do Gap in Childhood Pneumonia Assessment Among Frontline Health Workers in an Indian State .....	213
O 171 - Very Low Birth Weight Newborn Tracking Assessment: An Evaluation Evolving into Ownership Transfer .....	214
O 172 - Zika Outbreak in Brazil: Understanding Perceptions to Increase the Effectiveness of Preventive Action .....	215
O 174 - Development of National Evaluation Systems in Africa – A Comparative Study of Ghana and South Africa .....	216
O 175 - Moving from a Decentralised to a Centralised Corporate Evaluation System: Opportunities for Methodological Innovation Combining Contribution Analysis and Oecd/Dac Criteria .....	217
O 176 - A Methodological Framework for the Evaluation and Selection of Policies for Endogenous Regional Development. The Case of Greek Regions .....	218
O 178 - Evaluating the United Nations' Sustainable Development Goals: Dropout Prevention Program in Finnish Vocational Education .....	219
O 179 - Repurposing the Qsort Method and the Classic Simulation Method to Isolate Criteria That Underlie Programme Evaluability Decisions .....	220
O 180 - Design evaluation of the Master Plan for International Cooperation for Development .....	221
O 181 - Use of Conversational Mapping for Evaluative Processes .....	222
O 182 - The Changing Logic of the Northern Ireland Peace Programmes .....	223
O 183 - Evaluating Peace. Flexibility and Complexity in the Evaluation of the Norwegian Engagement in the Peace Process in Colombia .....	224
O 184 - Transitional Justice: Views from the Ground on How Afghanistan Fares .....	225
O 185 - Why Do Citizens not Participate in Community-based Monitoring? Insights from Uganda's Rural Water Sector .....	226
O 186 - Formative Participatory Evaluation Critical for Quality Assurance in African Universities .....	227
O 188 - Towards a General Theory of Monitoring and Evaluation: A Method to Balance Donor Budgets, Agency Capacity and Project Impact .....	228
O 189 - Examining the Extent and Effects of Disability-Sensitive Community Monitoring on Educational Service Delivery in Katakwi District, Uganda .....	229
O 190 - Integrating Social and Behavior Change Communication Indicators into a National Health Information Management System: The Ghana Experience .....	230
O 191 - Learning for Adapting: Lactational Amenorrhea Method (LAM) for Family Planning in the Philippines .....	231

O 192 - Improving Public Health Management Advances Resilient Societies: Evaluation Findings of a Workforce Capacity Building Program in Kenya and Bangladesh .....	232
O 193 - The Evolution of Spatial Operational Programmes' Evaluation in Greece: A Market Approach .....	233
O 194 - Making Evaluations Matter: Lessons Learnt from a Reorientation Process at the Media Development Organization DW Akademie .....	234
O 195 - Challenges Evaluating New Approaches to Programme Design: Evaluating Human-Centred Design for Behaviour Change in Adolescent Sexual and Reproductive Health .....	235
O 196 - Emergent Design of the Ecosystem Services for Poverty Alleviation (ESPA) End of Programme Review .....	236
O 197 - Economic Multiplier of an Asset Transfer Model: The Case of Heifer Malawi .....	237
O 198 - Quick Guide to Incorporate Young and Emerging Evaluators to Organizations: What Works? .....	238
O 199 - The EvalYouth National Chapters in LAC: a Comparative Study on the Strategies of Conformation and Sustainability of Voluntary Action .....	239
O 201 - Young and Emerging Evaluators in Latin America: Their Position and Role in the Evaluation Field .....	240
O 202 - A Scoping Review on the Current and Potential Use of Social Network Analysis for Evaluation Purposes .....	241
O 203 - Using Social Network Analysis to Evaluate the Network Formation Among Fellows in a South African Leadership Development Programme .....	242
O 204 - Through Local Eyes: Integrating Social Network Analysis and fieldwork to reconstruct local configurations in evaluations ..	243
O 205 - Toward more Resilient Project Evaluation: an Epic Battle between Professional and Political Logics .....	244
O 206 - How to Evaluate Complex Adaptive Systems? Case: Disaster Preparedness and Disaster Risk Reduction .....	245
O 207 - Systemic Evaluation in the Making: A Case Study .....	246
O 208 - How to Become a Resolute and Assertive Evaluator: The Role of Ethos and Fundamental Texts in the Education of Evaluators .....	247
O 209 - A Collaborative Implementation Evaluation among Governmental Instances in Brazil: An Innovative Experience for Health Care and Management Processes .....	248
O 210 - Examination of Evaluation Culture .....	249
O 211 - Newborn and Child Care Along with Related Health-Seeking in Bihar: An Exploration by Gender of the Child .....	250
O 212 - Dealing with Flexibility and Complexity: Lessons from Impact Assessment of the Shaurya Dal Programme of Government of Madhya Pradesh, India .....	251
O 214 - Evaluating Impact in Complex Environment – Dialogue on the Needs, Methods and Practices Based on Sitra's Impact Evaluations .....	252
O 215 - Using a Wide Lens to Take a Long View: How Systems Thinking Can Help Incorporate Long-Term Perspectives into Short-Term Evaluations .....	253
O 216 - Evaluating the Sustainability of Complex Interventions: The Case of Police Reform in DRC .....	254
O 217 - Managing Evaluation of Complex Interventions: Resilience in Evaluation .....	255
O 218 - Evaluative Resilience Through OODA Looping .....	256
O 219 - Evaluating Complex Systems for Strengthened Resilience: Experiences from the Africa Risk Capacity .....	257
O 220 - The Evaluation of Public Policies and the Politics of Evaluations: Views from Postmodernist and Complex Realities in Latin America .....	258
O 221 - Cultural Competency and Responsiveness in Evaluations: Promoting Resilience and Development in Challenging Contexts	259
O 222 - Skin in the Game: Epistemological, Political and Moral Contradictions In International Development Evaluation Discourse ..	260
O 223 - Small but Mighty? The Market for Evaluation in the Canadian Not-For-Profit Sector .....	261
O 224 - Evaluation of Financial Sustainability of Investment Facilities .....	262
O 225 - Great Expectations: An Enquiry into the Premise of Directing Institutional Investors Towards Emerging Markets, Especially in Times of Market Crises .....	263
O 226 - Method Study and Methods Engineering Techniques in Improving Business Processes and Outcome Driven Organizational Viability .....	264
O 227 - Impact Evaluation of Rural Finance Institution Building Programme on Food Security Status Among Beneficiaries in Anambra State, Nigeria .....	265
O 229 - Combining Realist Evaluation and Causal Loop Diagramming in Evaluating a Performance-Based Financing Intervention in the Ugandan Health Care Sector .....	266
O 230 - Assessing the Pertinence of Primary Healthcare Programmes' Objectives using a Hierarchical Fuzzy Approach: a case study in Brazil .....	267

O 231 - The Link between Evaluation Policy and Organisational Capacity for Evaluation: What Evaluation Scholars and Practitioners are Saying .....	268
O 232 - The Role of Evaluation for Supporting the Most Vulnerable in Kosovo .....	269
O 233 - The Role of Innovation in Supporting a Resilient Healthcare System in the UK; Can Evaluation Help (and, if So, How)? .....	270
O 234 - Post NPM: The Evaluation of Trust-Building Governance .....	271
O 235 - A Case Study of an Evaluation Association's Pathway to Professionalization: AEA's Past, Present, and Future .....	272
O 236 - Reinforcing the Link Between Learning and Action: Engaging Students in Grounded Learning About Evaluator Competencies and Evaluation Practice .....	273
O 237 - Strengthening Evaluative Thinking Skills for Achieving the Sdgs in the "Post-Truth Era" .....	274
O 238 - Development Evaluation Beyond Aid .....	275
O 239 - Stretching the Limits – Challenges and New Options in Applying the DAC Criteria to the Evaluation of Human Rights Interventions .....	276
O 240 - Mobile Data Collection is the Future for Evaluation but Often in Remote Poor Areas it is Often Fraught with Problems .....	277
O 242 - Evaluation Methods for the R&D Programs Dedicated to Solve Social Problems .....	278
O 243 - Evaluating Resilience as a Component of Transition (Development) Impact .....	279
O 244 - Evaluation Advisers Able to Embrace Resilience – Insights from Department of International Development (UK) Experience .....	280
O 245 - Building a Culture of Transparency: Managing Evaluation to Support Resilience .....	281
O 246 - Evaluations in Prisons – Reflections on Evaluation Practice in a Challenging Context .....	282
O 248 - Challenges of Evaluating Impact of Sustainable Food Security and Undernutrition Intervention: Case Study of Bukoba Rural, Republic of Tanzania .....	283
O 250 - Evaluators for Planetary Health .....	284
O 251 - Evaluation in Times of Crisis: Rethinking the Scope as well as the Methods in a Changing Environment .....	285
O 252 - Evaluating Partnerships in Research and Innovation within a Mission-Oriented Policy Framework .....	286
O 253 - Impact Evaluation of ACCRA Ethiopia Programme as a Moment of Triple Loop Learning – a Critical Contribution in Turbulent Times? .....	287
O 254 - Building Livelihood Resilience in Rural Malawi: Constraints to Credible Impact Evaluation and Synthesis .....	288
O 255 - Taking Impacts to Mixed Methods Evaluations in Ex-Post and Ex-Ante Settings .....	289
O 256 - Counter-factual Impact Evaluation of EUREKA Network Projects and Cluster Projects – Strengthening R&I Policy Resilience through a Multi-Methodological Approach .....	290
O 257 - A Perfect Match: A Demonstration of Pattern-Matching Techniques to Test Complex Program Theories .....	291
O 258 - Applying the Principles of Outcome Harvesting for Monitoring and Evaluating Development Interventions .....	292
O 259 - Applying Outcome Evidencing to Unpack Complexity and Contextualise Results in a Livestock Project in Zimbabwe .....	293
O 260 - Play It Different This Time, Sam! Involving Teachers and Educators in Evaluation of Drop-Out Prevention Programs in Schools .....	294
O 261 - Sustainable Development in The Context of Vocational Competencies and the Goals of Agenda 2030 .....	295
O 262 - How Can We Use Evidence Better for Achieving Social Justice? Lessons from Oxfam's Evidence for Influencing Conference .....	296
O 263 - Evaluating Unintended Effects and Adaptation in Countries Affected by Fragility, Conflict and Violence .....	297
O 264 - Values, Judgment and Reasoning in an Era of Evidence .....	298
O 265 - Engaging with the Affected Population in the Evaluation of Humanitarian Responses .....	299
O 266 - Use of the Evaluation for the Programme on Violence Against Children in School in Jordan .....	301
O 267 - Evaluation of Environmental Protection Processes. Application in a Major Industrial Installation in Greece .....	302
O 268 - Becoming Bigger, Better, Smarter: A Review of the Evaluability of Climate Change Proposals Presented to the GCF .....	303
O 269 - Re-Thinking Hierarchical Evaluation Systems Towards Evaluation Theory of Change that Defines Sustainability in African: Using the Psychological Perspective to Demonstrate .....	304
O 270 - More Often than Not, Contribution and Attribution are Confused In Impact Evaluation. This Paper Splits These Elements Clearly .....	305
O 271 - Addressing Unintended Effects in Development Aid Evaluations .....	306
O 272 - Mixed Method Strategies for Strengthening Internal and External Validity: Combining QCA with QEM to evaluate household vulnerability in Uganda .....	307
O 273 - National Evaluation Policies in Europe: Observations and Learning from the VNRs .....	308
O 274 - Inclusive System Evaluation: Gender Equality, Environments and Marginalized Voices: A UN Women Guide for Empowering Voices in the SDG Era .....	309

<b>O 275 - New Indicators for Sustainable Development Goals Targets; Technical Review of the Voices of the Hungry Project's Food Insecurity Experience Scale</b> .....	310
<b>O 276 - Evaluation as a Live Practice – in the Space Between Audit and Learning</b> .....	311
<b>O 277 - Integrating Gender in Evaluation in the Context of a Multilateral Development Bank</b> .....	312
<b>O 278 - Evaluation Function in the Spanish Cooperation: A Changes Proof System?</b> .....	313
<b>O 279 - Performance Leadership: The Key to Strengthening Evaluation Systems and Program Outcomes</b> .....	314
<b>O 280 - Evaluating “Leadership”: Thinking Modular in an Organization-Wide Evaluation</b> .....	315
<b>O 281 - Examining Personal and Organisational Impacts of a Leadership Intervention on a Non-Governmental Organisation</b> .....	316
<b>O 282 - Decrypting Contextual Data to Reduce Bias: Case Study in Women's Rights Movement from Non-Western Lens</b> .....	317
<b>O 284 - Considering Policy Environment in Providing Policy Advice – Evaluation of Knowledge Sharing Program (KSP) as Foreign Aid</b> .....	318
<b>O 285 - Building Resilience to Recurrent Refugee Crisis: an Evaluation of the Uganda Comprehensive Refugee Programme</b> .....	319
<b>O 286 - Whose Resilience? Considering Gender and Other Systemic Power Dimensions when Assessing Resilience Capacities through Impact Evaluations</b> .....	320
<b>O 287 - Consulting as a Second Act Career: Transition Issues</b> .....	321
<b>O 288 - Evaluators in the Gig Economy</b> .....	322
<b>O 289 - When Power and Money meet Evaluation – Revisiting the “Matrix of Values”</b> .....	323
<b>O 290 - Adaptive and Robust Evaluative Reasoning in Dynamic Evaluation Contexts</b> .....	324
<b>O 291 - How Does Evaluation Reshape Accountability Systems? Insights from a Case Study on Higher Education Reform in Italy</b> ..	325
<b>O 292 - Exploring the (Un)Intentional Consequences of Teacher Education Programming in Uganda, Africa</b> .....	326
<b>O 293 - Evaluation of a Low Cost Mobile Device Initiative for First Year University Students</b> .....	327
<b>O 294 - Recognizing and Acting on Several Levels of Evaluation Use: A Synthesis of Existing Models</b> .....	328
<b>O 295 - Evaluation Use and Resilience: Going Beyond Interventions</b> .....	329
<b>O 298 - Ex-post Impact Evaluation for Public-Private Partnership Projects Using Multiple Stakeholder Perspectives</b> .....	330
<b>O 299 - How to Tackle Attribution in Public-Private Partnerships?</b> .....	331
<b>O 300 - Evaluation Challenges and Opportunities within Corporate Workplace Sustainability Programming: Integration of SDG 8 Economic Growth and Decent Work</b> .....	332
<b>O 301 - Critical Analysis of the Challenges and Opportunities for the Evaluation in the Public Sector in Uganda</b> .....	333
<b>O 302 - The Evaluator's Dilemma: Ethical Issues in Evaluations with Multiple Stakeholders with Different Agenda</b> .....	334
<b>O 303 - For Better Evaluation Capacity and a Stronger Evaluation Profession in Portugal: Challenges and Opportunities</b> .....	335
<b>O 304 - When Social Impact and Developmental Evaluation Meet in Japan – Professional Development of a New Breed of Evaluators</b> .....	336
<b>O 305 - A Survey on Organizational Readiness for Evaluation in Turkey</b> .....	337
<b>O 306 - Discovering Behavioural Patterns of Major Transport Projects: Lessons from an Ex-Post Evaluation</b> .....	338
<b>O 307 - INTERREG Evaluations – The Challenge of Demonstrating Results and Value of Territorial Cooperation</b> .....	339
<b>O 308 - Evaluating Job Quality in the EU Funds and Programmes that Support Smart, Sustainable, and Inclusive Growth</b> .....	340
<b>O 309 - Evaluation Of The European Maritime And Fisheries Fund, Balancing Between Scarce Resources And Broad Thematic Scope</b> .....	341
<b>O 310 - Evaluation Studies Supporting Complex Programs in the Field of Culture</b> .....	342
<b>O 311 - Quasi-Experimental Impact Assessment: Evaluating Urban Centers of Culture, Art, Science and Sports in Brazil</b> .....	343
<b>O 312 - Building the Empirical Evidence of Youth Participatory and Collaborative Evaluation</b> .....	344
<b>O 313 - Capturing the Voices Of Marginalised People: Designing, Commissioning and Delivering Disability Inclusive Evaluations</b> ..	345
<b>O 314 - Conducting Gender-Sensitive Needs Assessment. The Case of Caregivers Needs</b> .....	346
<b>O 315 - Multiple Outcomes of Building a Knowledge Management System to Manage Reflexive Impact Evaluations for Agricultural Research</b> .....	347
<b>O 316 - The Missing Diagnosis and the Evidence Trap: A Tale of Inadequate Program Design</b> .....	348
<b>O 317 - What Can the Evaluation Community Learn from the Practice of Literature Reviews? Lessons Learnt from Eight Case Studies</b>	349
<b>O 318 - The African Countries on a Path to Increase the Use of Evidence for Policymaking</b> .....	350
<b>O 319 - Evaluating Monitoring, Evaluation, Accountability and Learning (MEAL) capacity: lessons learned from Save the Children</b>	351
<b>O 320 - Applying the Peer Review Method to Civil Society Evaluations</b> .....	352

O 321 - Redesigning the Evaluation Architecture for More Resilient Evaluations; The Case of Integrated Care in the UK .....	353
O 322 - Designing Evaluations to Provide Evidence to Inform Action in New Settings .....	354
O 323 - Lessons Learnt & Best Practices Managing Prospective Evaluations in a Constantly Changing Environment: Studies from Multi-Country Implementation in Gavi-Supported Countries .....	355
O 324 - Evaluation of Inter-linkages and Trade-offs Between the Sustainable Development Goals .....	357
O 325 - 10 Years of Monitoring Progress towards Sustainable Development in the EU – What Lessons can be Learned for the SDGs? .....	358
O 326 - Socio-Economic Impacts from Innovation Scale-Up: A Framework for Evaluating Urban Innovation Programmes .....	359
O 327 - Evaluation of EU Assistance to Ukraine: How to Ensure Gender Balance .....	360
O 328 - Rethinking Evaluation Methods, Design and Criteria for Assessing Policy and Programming Coherence when Evaluating Diplomacy, Trade and International Assistance Interventions .....	361
O 329 - The Role of Evaluation Agency for The Indonesian Government Policy in Water Access .....	362
O 330 - New Solutions for Europe ...and Beyond? The Role of Africa in Leading a Revision of the Realist Evaluation Lens .....	363
O 331 - The Rise of Realist Evaluations in Randomised Control Trials of Employment-related Interventions in the UK .....	364
O 332 - Addressing Complexity in Programme Theory: Experiences in Applying the Actor-Based Change (ABC) Framework to International Development Programmes and Evaluations .....	365
O 333 - The Roles and Sustainability Measures of Evaluation Networks and Associations in the Professionalization of YEEs in Developing Countries .....	366
O 334 - The Role of Voluntary Organizations for Professional Evaluations on Promoting Evaluation Culture: Experience and Challenges in Tanzania .....	367
O 336 - Introducing a Monitoring & Evaluation System for the Russian Federation's International Development Assistance Program – A Look Behind the Curtain .....	368
O 337 - No Pain No Gain: Controversies and Complexity in Modeling Mental Health Care Centers in the Amazon Region, Brazil ...	369
O 338 - Towards a General Theory of Monitoring and Evaluation: A Standard for the Scope of the Evaluation in Terms of Reference .....	371
O 339 - Sexual and Reproductive Health Challenges of Adolescent Females in Some Communities of Akwa Ibom, Nigeria .....	372
O 340 - Evaluating Digital Tools for Behaviour Change using Big Data .....	373
O 341 - Prospects for Democratic Evaluation in a Polarised and Mediatized Society .....	374
O 342 - Trump Our Best Advocate for Evaluation !? .....	375
O 343 - Deliberative Liberalism as a Strategic Approach for Evaluation in Turbulent Times .....	377
O 344 - Help! I'm Doing An Impact Evaluation, What Evidence Do I Need? .....	378
O 346 - Using Social Media for Programme Theory Development: The Role of Emerging Technology in Complex Programme Evaluation .....	379
O 347 - Key Challenges and Lessons of Making the Best Use of Theories of Change in Thematic Evaluations .....	380
O 348 - Evaluation As A Social Practice .....	381
O 349 - The Reflective Practitioner – The Road To Innovation .....	382
O 350 - Social Return on Investment (SROI): Redefining Value and Expanding Possibilities .....	383
O 351 - Bringing Evaluation Tools into Impact Investing: Using Theories of Change to Understand the Various Financial Mechanisms Used by Impact Investors .....	384
O 352 - Evaluating Corporate Responsibility in Research and Innovation .....	385
O 353 - Using Social Return on Investment Analysis to Assist Companies in Optimizing Their Social Value Towards the SDGs .....	386
O 354 - Five, Six or Seven Waves of Evaluation? Across the Western World Since 1965–2018 .....	388
O 355 - Developing the Field of Evaluation to Promote Resilience and Action in Critical Times: Values in Evaluation: Philosophical, Ethical and Theoretical Considerations .....	389
O 356 - The Tragedy of Evaluation .....	390
O 357 - Exploiting “Big Data” to Assess Vulnerability and Resilience: Natural Experiments and Observational Research Designs ...	391
O 358 - School-based Data Collection in the Context of an Earthquake: Evaluation of a Comprehensive Sexuality Education Program in Mexico City .....	392
O 359 - Methodological Decisions for Evaluating Rural Community Resilience Investments in Malawi: Maximising Rigour and Utilization .....	394
O 360 - Listening to What Children Need, Think and Recommend .....	395
O 361 - Evaluation and Statistics: The New Partnership .....	396
O 362 - Integrating Qualitative Research Methods into Economic Evaluation .....	398

O 363 - Visualizing Change: Mixed Method Data Analysis in Evaluation with ImpactMapper .....	399
O 364 - Cultural Development and the Overcome of Poverty: Evaluation of the 'Servicio País Cultura' Program in Valparaíso, Chile	400
O 366 - Evaluation of the Growth and Poverty Reduction Strategy Paper III of Cabo Verde .....	401
O 367 - Sustainable Development through Alternative Energy: The Implication of Socio-Economic Development Programmes .....	402
O 368 - Analyzing the Program Theories of the Policies Against Poverty and Social Exclusion. The "Income of Inclusion" in Italy. ....	403
O 369 - Evaluating Market System Development Programmes: The Value of Sector Knowledge .....	404
O 370 - The Role of Parliamentarian Evaluation Forum for Creating Resilient Society .....	405
O 371 - Evaluation and New Public Management: Substitution or Complementarity? .....	406
O 372 - Institutionalizing the Evaluation as an Improvement Mechanism for Public Interventions .....	407
O 373 - Sustaining Monitoring and Evaluation in Health Care .....	408
O 374 - Evaluation Standards Development and Use: The case of the International Development Evaluation Association (IDEAS) and UNICEF .....	409
O 375 - Evolution of the Canadian Evaluation Society's Professional Designation, 2010–2018: Implications for Resiliency of the Field	411
O 376 - Evaluation as an Agent of Change .....	412
O 377 - Constructing the Outcome Monitoring System for Portfolio of Biotechnology Programs with Mixed Methods and Benefits Diffusion Model .....	413
O 378 - Getting it Right: A Mix of Evaluation Methodologies for Complex Research for Development Programs .....	414
O 379 - The Use of Social Media in Monitoring and Evaluation: Opportunities and Challenges .....	415
O 380 - Benefits of Using Big Data in Evaluations and How to Overcome Technical Constraints .....	416
O 381 - The Best of Both Worlds? Integrating Traditional Evaluation and Data Science Methods .....	417
O 382 - Redefining Monitoring and Evaluation Systems for Development effectiveness of Watershed Development Programs .....	418
O 383 - A Realist Evaluation Approach to Studying Impacts of Dance Based Interventions on Parkinson's Patients .....	419
O 384 - In Search for a Realist "Fiat": Lessons from Operating Realist Evaluation with Limited Budgets .....	420
O 385 - Combining Realist with Developmental and Utilization-Focused Evaluation Methodology: Conflict or Symbiosis? .....	421
O 386 - Although Realistic Evaluation has Improved Ascertaining Causalities, it is Still Fraught with a Number of Methodological Flaws Worth Reviewing .....	422
O 387 - Revisiting Theory of Change of Interventions in Complex Systems .....	423
O 388 - Deepening Participatory Practice to Build Inclusive Resilience; Where is the Evidence? .....	424
O 390 - Fostering Resilience through "Strategic" Monitoring and Evaluation: Evidence from the National Science Foundation Efforts to Build/Diversify the Scientific Workforce .....	425
O 391 - Comparing the Institutionalization Approaches of Korea and Netherlands Regarding the Integration of Foresight, Technology Assessment, and Evaluation .....	426
O 392 - Evaluating Effects of Public Support to Business R&D and the Economic Crisis: Spanish Evidence .....	428
O 393 - The Role of Evaluation in Helping Build Resilience to Violent External Shocks: Lessons Learned and the Way Forward .....	429
O 394 - "You Can't Get There from Here": A Program Evaluation in Iraq .....	430
O 395 - Making Resilience Visible: Results from a Rigorous Impact Evaluation of a Land-Use Planning Intervention in the Philippines	431
O 396 - Organisational Reviews of Civil Society Organisations: How Can Usefulness Improve While Costs Are Reduced? .....	432
O 397 - Relevancy Evaluation in Cooperatives' Performance. A Guiding Material for Cooperatives' Performance Evaluation .....	433
O 398 - Evaluating the Work of German 'Stiftungen' Through Bonding, Bridging and Linking Social Capital: A Case-Study on Kyrgyzstan .....	434
O 400 - Multi-Stakeholder Initiatives – the Case of EITI: Paving Pathways to Impacts through Evaluations. ....	435
O 401 - Engaging The Private Sector for Agricultural Development – an Evaluation of Germany's Policy Approach .....	436
O 403 - Exploring Methodologies to Measure Power Transformation in Social Fabrics .....	437
O 404 - From Impact Heterogeneity to Implementation Heterogeneity: Equity Measurements for Informed Programming .....	438
O 405 - Evaluation Methods Old and New! The Use of Bibliometrics and Expert Panel Review to Better Evaluate Scientific Impact ..	439
O 406 - Ex-Ante Evaluation of Research Infrastructures: Benefits, Methods and Approaches, Steps in Implementation, and Other ...	440
O 407 - The Socioeconomic Impact Assessment of Research Infrastructures: A Critical Review of Existing Approaches .....	441
O 408 - From Theory-based Evaluations to Theory-based Monitoring-and-Evaluations? .....	442
O 409 - Graduate Tracking System as a Measure of Evaluation of the Educational System .....	443

O 410 - An Innovative Evaluation Approach Towards High Quality and Rigorous Evidence: Documentation of Program Processes . . .	444
O 411 - A Stakeholder Centric Approach for Measurement Learning and Evaluation in Complex Interventions: The Complexity Responsive Stakeholder Centric Evaluation Framework . . . . .	445
O 413 - Evaluating SDGs in No One Left Behind Lens . . . . .	447
O 414 - The Importance of the 'S' in SDG: Lessons from a WASH Project in Ghana and Kenya . . . . .	448
O 415 - Citizen Participation and Inclusive Governance for Sustainable Development in Uganda: Lessons from Parliamentary Outreach Programs . . . . .	449
O 416 - Sustainability Report of Municipality of Agios Dimitrios, Greece The society evaluates the local authorities . . . . .	450
O 417 - How a VOPE comes true: RedEvalCR . . . . .	451

## Panels

452

The Role of Evaluation in the Face of Forced Displacement and the Emerging Humanitarian / Development Nexus . . . . .	453
PC 001 - The Importance of the Humanitarian/Development Nexus in Realizing the SDGs – Challenges for Evaluation . . . . .	454
PC 002 - Impact Investing and Humanitarian Aid: Can Value for Money Approaches Help Addressing the Gap? . . . . .	454
Stories and Scars From the Field: Guidelines for Managing Conflict in the Conduct of Evaluation . . . . .	456
PC 004 - Conflict within Evaluation of the Improving Schools Program . . . . .	457
PC 005 - Conflict within Evaluation of the Managing for Development Results Strategy . . . . .	457
PC 006 - Conflict within Evaluation of the Western Australian Tough on Graffiti Strategy . . . . .	457
PC 007 - Conflict within Evaluation of the Improving Schools Program . . . . .	457
PC 008 - Reviewing Conflict Across the Case Studies and the Literature . . . . .	457
There is no Resilience without Equity: When will Our Profession Finally Act to Reverse Global South Asymmetries in Global Evaluation? . . . . .	458
PC 009 - Hidden Power and Privilege in Evaluation: A Latin American Perspective . . . . .	458
PC 010 - South Asian Evaluators Tackle Power Asymmetries in South Asia . . . . .	458
PC 011 - Made in Africa Evaluation: Why Africa-Centric Evaluation Matters in Global Evaluation . . . . .	459
PC 012 - Funders Can Do Better and They Know It: What Is Stopping Them? . . . . .	459
PC 285 - Who holds the Power in the Evaluation System? Perspectives from the Global South . . . . .	459
Evaluation for the Anthropocene: A Call to Action for a Sustainability-Ready Evaluation . . . . .	460
PC 014 - Panel Contribution . . . . .	460
PC 015 - Panel Contribution . . . . .	460
PC 016 - Panel Contribution . . . . .	460
Leveraging Longitudinal Data for Resilience Measurement . . . . .	461
PC 017 - Insights and Lessons Learned From High Frequency Recurrent Monitoring in Ethiopia . . . . .	461
PC 018 - Two Years Later: Determinants of Coping and Recovery to the Gorkha Earthquake in Nepal . . . . .	462
How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals and International NGO Evaluators . . . . .	463
PC 019 - How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals and International NGO Evaluators: Views from the Chair . . . . .	464
PC 020 - How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals and International NGO Evaluators: Views from NGOs and Participants . . . . .	464
PC 021 - How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals, International NGO Evaluators: A View from ADB & AusAid . . . . .	464
PC 022 - How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals and International NGO Evaluators: the World Bank/IEG & Maastricht . . . . .	465
PC 023 - How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals and International NGO Evaluators: A View from Development Agencies . . . . .	465
PC 024 - How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals and International NGO Evaluators: A Bilateral and Academic View . . . . .	465
Tackling Complexity Theory in Evaluation: Methodological Challenges, Lessons Learned and Practical Solutions . . . . .	466

PC 025 - Tackling Complexity: Gender Mainstreaming as a Roadmap .....	467
PC 026 - Complexity-Sensitive Evaluation: A Five Step Approach .....	467
PC 027 - SDG-Responsive Evaluation: Are We Getting Better at Managing Complexity? .....	467
PC 028 - Striking a Balance Between Theory, Methodological Concerns and Practical Solutions for Evaluation? .....	468
PC 029 - Implications of Complexity Theory for Evaluation .....	468
How Can Evaluation Help Bureaucracies be, Well, Less Bureaucratic? .....	469
PC 030 - Why Interest in Bureaucracy and Potential Roles for Evaluation .....	470
PC 031 - The Disconnect Between Outcome-Oriented Public Services and Activity-Based Bureaucratic Practices .....	470
PC 032 - Management Paradigms, Accountability, Evaluation, and Bureaucracy in (Un)Predictable Contexts. ....	470
PC 033 - Evaluation Quality Assessment: Useful Tools vs. Bureaucratic Impediments to Learning? .....	470
PC 034 - Panel Contribution .....	471
PC 035 - Evaluative Thinking and Bureaucracy .....	471
Chairing the Session on “Building Strong Monitoring and Evaluation Systems for Global Partnership Programs” .....	472
PC 036 - Independent Evaluation Group .....	472
PC 037 - Global Partnership for Education (GPE) Experience .....	473
PC 038 - Global Environment Facility (GEF) Experience. ....	473
The Evaluation Marketplace: Exploring the Market Dynamics of Evaluation .....	474
PC 039 - The Evaluation Marketplace – A Conceptual Framework .....	474
PC 040 - The Evolving Market for Systematic Evaluation in Canada .....	474
PC 041 - Does Reputational Capital Matter within Italy’s Quasi-Market for Evaluation? .....	475
PC 042 - The US Federal Evaluation Market .....	475
PC 043 - The Commercial of Evaluation: A Danish Case Study .....	475
Evaluation for Improving Public Sector Efficiency .....	476
PC 044 - Public Management by Objectives. ....	477
PC 045 - How to Measure Good Governance: The Twelve European Principles of Good Democratic Governance .....	477
PC 046 - Panel Contribution .....	477
Powers and Privileges: The Non-Visible Side of Evaluations. Lessons From Latin America. ....	478
PC 050 - Young and Emerging Evaluators (YEEs): Experiences, Challenges and Opportunities in the Face of Inclusion with Equity ..	479
PC 051 - Power and Gender Issues: Reflections and Proposals From the Field of Evaluation .....	479
PC 052 - New Approaches to Evaluate from the Feminist Perspective: Towards a Non-Hegemonic Masculinity in Evaluation .....	480
Measuring the Impact of Structural Reforms carried out in Times of Crisis .....	481
PC 053 - Panel Contribution .....	482
PC 054 - Panel Contribution .....	482
PC 055 - Panel Contribution .....	482
PC 056 - Panel Contribution .....	482
Looking to the Criticality of Oversight and the Necessity of Working Together. ....	483
PC 058 - Audit and Evaluation: Working Collaboratively to Support Accountability .....	484
PC 060 - Evaluation ‘Architectures’ – Implications of How we Work Together. ....	484
Privacy, GDPR and the Responsible Use of Data in Evaluation Systems. ....	485
PC 061 - Designing Monitoring and Evaluation Systems for Resilient Societies .....	485
Outcome Harvesting Customized for Monitoring and Evaluation in the Dutch Dialogue & Dissent Lobbying and Advocacy Programme for Development .....	487
PC 062 - The Challenges of the use of Outcome Harvesting for Consortia in the Dutch Dialogue and Dissent (D&D) Development Program .....	488
PC 063 - Using Outcome Harvesting as an Approach for Monitoring Lobby & Advocacy Programs .....	488
PC 064 - Outcome Harvesting: Oxfam Novib’s Large-Scale Deductive Analysis of Outcome Statements Validating a Theory of Change .....	488

<b>PC 065 - When Performance Evaluation Using DAC Criteria won't do: Findings from an Outcome Harvesting mid-term Evaluation of four NIMD Country Programs</b> .....	489
<b>Great Expectations: An Enquiry into the Premise of Directing Institutional Investors to Emerging Markets, Especially in Times of Market Crises</b> .....	490
PC 066 - Panel Contribution .....	491
PC 067 - Panel Contribution .....	491
PC 068 - Panel Contribution .....	491
<b>Evaluation Professionalisation in Practice: What are the Next Steps to Follow?</b> .....	492
PC 069 - The Case for Professional Designation in Evaluation .....	493
PC 070 - Do Evaluators Need a Unique Set of Skills? .....	494
PC 071 - Three Levels of Professionalization .....	495
PC 072 - Young Emerging Evaluators: The Role of Higher Education in Building Their Capacity .....	495
<b>Evaluation for a Caring Society. The Potential of Care Ethics for Evaluation</b> .....	497
PC 073 - Democratic Caring Evaluation .....	498
PC 074 - Democratic Evaluation and Care Ethics .....	498
PC 075 - Democratic Evaluation and Care Ethics .....	498
<b>Evaluating Value for Money in Complex, Adaptive Development Programmes</b> .....	499
PC 076 - Panel Contribution .....	499
PC 077 - Panel Contribution .....	500
PC 078 - Panel Contribution .....	500
<b>Evaluations to Transform: Addressing Gender Imbalances in Culturally Diverse Contexts</b> .....	501
PC 079 - South Asia Case Study Development for Gender Transformative Approach to Evaluation .....	502
PC 081 - African Wide Experiences of Gender Transformative Approach in Evaluation .....	502
PC 082 - Evaluations to Transform: What Competencies Do We Need to Address Gender Power Issues in Diverse Cultural Contexts? Lessons from Latin America .....	502
<b>Spreading the Good News: Applications of Theory-Based Causal Analysis in International Development</b> .....	504
PC 083 - Three Case-Based Approaches in One: Combining Pattern Matching, Process Tracing and QCA Under Real-World Evaluation Conditions .....	505
PC 084 - How to Increase Internal Validity in Theory-Based Evaluation? Integrating Comparative Case Studies and Theory-Testing Process Tracing .....	505
PC 085 - Panel Contribution .....	506
<b>Conducting Impact Evaluation in Difficult Contexts Using Geo-spatial Analysis</b> .....	507
PC 086 - Conducting Impact Evaluation in Difficult Contexts Using Geo-spatial Analysis .....	507
<b>Building European Evaluation Capacity: The Role of Young and Emerging Evaluators</b> .....	509
PC 087 - Why the Future of Evaluation Rests on Improving What Novice Evaluators Know and Can Do .....	509
PC 088 - Creating YEE initiative in ECA region (Eastern Europe, Central Asia & South Caucasus): Opportunities and Challenges .....	510
PC 089 - A Comprehensive Map of European Evaluation Capabilities .....	510
PC 090 - Is Professionalization of Evaluation a Global Movement? Can It Be? Ought It Be? .....	510
PC 091 - Discussant .....	510
<b>From Measuring to Managing for Results: Building in-country Monitoring &amp; Evaluation (M&amp;E) Capacities and Systems</b> .....	511
PC 092 - Prime: How to Build and Implement a Global Training on M&E in Rural Sector .....	512
PC 093 - PRiME: Establishing a Global Certification Framework on M&E in Rural Development .....	512
PC 094 - Linking Knowledge Management and Assessments of Capacities of In-Country M&E Systems .....	512
PC 095 - Results from Conducting In-Country M&E Capacity Assessments .....	513
PC 096 - Linking IFAD's Results Agenda with Strengthened In-Country M&E Capacities .....	513
PC 097 - IFAD's Approach to Strengthen Results-Based Management Systems in Partner Countries .....	514
<b>What Makes a Good Policy? WFP Lessons on Policy Formulation and Practicalities</b> .....	515
PC 098 - Panel Contribution .....	515
PC 099 - Panel Contribution .....	516

PC 100 - Panel Contribution .....	516
PC 101 - Panel Contribution .....	516
<b>Why Evaluation Theory Should be Used to Inform Evaluation Policies: Reflections on the Theory-Policy-Practice Connection .....</b>	<b>517</b>
PC 102 - The Evaluation Theory-Policy-Practice Relationship .....	517
PC 103 - Practicing What We Preach – Making the Case for a Stronger Integration of Evaluation Theory in Evaluation Policies .....	517
PC 104 - Six Characterizations in Search of an Evaluation Theory-Policy- Practice Relationship.....	518
PC 105 - A Behavioral Design for Evaluation Policy Reform in Italy.....	518
PC 106 - Discussant .....	518
<b>How Much Rigor is Enough? Impact Assessment for Better Decision Making and Increased Resilience .....</b>	<b>519</b>
PC 107 - Comparative Analysis of Different Levels of Rigor when Measuring Social Return on Investment (SROI) .....	520
PC 108 - Broadening Boundaries to Improve Estimates of Program Impact .....	520
PC 109 - Rigour Used by CBA in the EU Cohesion Policy Context and Stories About How is CBA Influencing (or not) Decision-Making.....	520
PC 110 - Triple Bottom Line: Is the Perfect Enemy of the Good? .....	520
PC 111 - Apples and Oranges? Thoughts on Approaches of Impact Investing Community and MDB Private Sector Operations .....	521
PC 112 - The Adoption of Social Risks Framework to Support Social Return on Investment Analysis: a Client Testimony.....	521
<b>The Voluntary Evaluator Peer Review Model: Building Professional Accountability Through Professional Development .....</b>	<b>523</b>
PC 113 - Pros and Cons of an Auditing Approach to Capabilities Assessment – Reviewee Perspective .....	523
PC 114 - Pros and Cons of a Peer Review Approach to Capabilities Assessment – Reviewee Perspective .....	524
PC 115 - Pros and Cons of Each Approach – Reviewer Perspective .....	524
PC 116 - Transportability of VEPR Across Cultures .....	524
<b>New Evaluation Approaches for Changing Feminist Times .....</b>	<b>525</b>
PC 117 - Inclusive Systemic Evaluation for Gender equality, Environments and Marginalized voices (ISE4GEMs): A New Approach for the SDG Era .....	525
PC 118 - Democratic Evaluation and Democracy: Exploring the Reality .....	526
PC 119 - New Evaluation Approaches for Changing Feminist Times: What Evaluation Agenda for the Fourth Feminist Wave?.....	526
<b>Adapting Outcome Harvesting for Monitoring and Learning (M&amp;E).....</b>	<b>527</b>
PC 120 - Applications and use of Outcome Harvesting to Strengthen Monitoring, Evaluation and Learning for Canadian Red Cross projects.....	528
PC 121 - The Outcome Harvesting Learning Journey of AADK – From Undertaking an External Evaluation to Building an OH Community in Denmark .....	528
PC 122 - Harvesting Outcomes in a Workshop: Ideas and Practice from a Cross-Border Network of Civil Society Organisations .....	528
PC 123 - Using Outcome Harvesting for Adaptive Management in Human Rights and Governance Programming .....	529
<b>Using Innovative Geospatial Analysis to Evaluate Geographic Footprint of Development Interventions.....</b>	<b>530</b>
PC 124 - Putting Your Money where Your Mouth is: Geographic Targeting of World Bank Projects to the Bottom 40 Percent .....	531
PC 125 - Geospatial Analysis of World Bank Projects in Mexico.....	531
<b>Assessment Methodologies for Transformational Governance and Community Resilience: Applications in Gang Violence Prevention, Countering Violent Extremism and Resilient Governance Systems .....</b>	<b>532</b>
PC 126 - Measuring Individual and Community Resilience: Applications to Gang Violence in Central America and CVE in Tunisia ..	533
PC 127 - Fragility Resilience Assessment Method (FRAMe): An Approach to Designing and Assessing Transformational Governance Resilience Programs .....	533
<b>The Economics of Resilience Returns: When 1 + 1 is Greater Than 2.....</b>	<b>535</b>
PC 128 - Evaluating the Returns to Community-Based Climate Resilience Interventions .....	535
PC 129 - Rising Every Time We Fall: The Economics of Resilience .....	535
<b>Using Monitoring and Evaluation to Help Societies Become More Resilient .....</b>	<b>536</b>
PC 130 - The Role of Monitoring and Evaluation in DG ECHO's Interventions Relating to Resilience Building and Disaster Response ..	536
PC 131 - Challenges to Evidence Quality in Evaluations of Humanitarian Action .....	537
PC 132 - Can Monitoring Help European Societies to Become More Resilient to Migratory Movements and Protect the Most Vulnerable? .....	537
<b>Addressing Ethics and Values in Evaluation.....</b>	<b>538</b>

PC 133 - Addressing the Difficulty for Value-Oriented Evaluations: Could a Moral Social Science Offer an Ethically Relevant Perspective? .....	538
PC 134 - Valuing Without Shared Values .....	539
PC 135 - Ethical Justifications for Collaborative Approaches to Evaluation .....	539
PC 136 - Valuing in the Service of the public Interest(s) in Our Complex World .....	539
PC 137 - The West and the Rest: an Exploration of the Role of Societal Values in Shaping Evaluation for Development .....	540
PC 138 - The Normative Political Characteristics of Professional Ethics in Evaluation: the Case for Democratic Professionalism .....	540
Learning through Principles Discovery Methods when the Program Theory is Unclear: A Data-Driven Approach to Understanding Context-Mechanisms-Outcomes Associated with Impacts .....	541
PC 139 - Applications of Principled Discovery Methods when the Theory of Change of Interventions is Unclear .....	542
PC 140 - Understanding and Identifying Relevant Contexts in Maternal Health Settings in India .....	542
PC 141 - Applying Principled Discovery Approaches to Evaluating the Impacts of a Dance-based Intervention Focused on Clients with Parkinson's Disease .....	543
PC 142 - Taking Intersectionalities Seriously: The Role of the Importance of Identifying Multiple Intersecting Contexts in Addressing Inequities .....	543
Privacy by Design: Principles to Balance Impact Measurement with Privacy, Security and Safeguarding .....	544
PC 143 - Live Data in Impact Assessment: Privacy and Security Considerations .....	544
PC 144 - Putting Privacy, Safeguarding and Gender at the Heart of Evaluation Design and Digital Data Processing .....	545
From Theory to Real World Evaluation: How Could Resilience Help Fill the Gap? .....	546
PC 150 - From Theory to Real World Evaluation Capacity Development: A Transformative and Resilient Perspective from West and Central Africa .....	547
PC 151 - From Theory to Real World Evaluation: How Could Resilience Help Fill the Gap? Independent Evaluation, DV and Real Participation .....	547
PC 152 - From Theory to Real World Evaluation: How Could Resilience Help Fill the Gap? Evaluation Uses / Users and Gender Gaps .....	548
Evaluating Complex Governance Interventions: Methodological Approaches to Analyse the Effects of Budget Support Programmes .....	549
PC 153 - A Mechanism-Centred Approach to Evaluating Complex Aid Interventions: The Case of Accompanying Measures to General Budget Support .....	550
PC 154 - Evaluating the Exit from Aid: Effects of the Exit from General Budget Support .....	550
PC 155 - Evaluating Budget Support Operations: A Comparison of the IEG and OECD-DAC Approaches and some Lessons .....	551
Evaluation Synthesis: Enhancing the Use Of Evaluation Findings for Decision Making and Reform Processes – Experience from IFAD And WFP .....	552
PC 286 - Panel Contribution .....	553
PC 156 - Panel Contribution .....	553
PC 157 - Panel Contribution .....	553
PC 158 - Panel Contribution .....	554
Meeting of Thematic Working Group for Emerging Evaluators .....	555
Designing and Implementing Adaptive Evaluations for Complex Programmes in Conflict Zones .....	556
PC 159 - Adapting to Adaptive Programming: Rethinking Roles, Relations and Relevance of Evaluation In a Complex Accountability Programme .....	557
PC 160 - Contribution Analysis and Conflict Zones: Assessing Impact in Afghanistan .....	557
PC 161 - Measuring Moving Targets: Evaluating Adaptive Market Development Programmes in Fragile Contexts .....	557
Evaluation Standards and Competencies for Resilient Societies .....	558
PC 162 - The Role and Future of Standards in North America .....	559
PC 163 - Evaluation Standards and Competencies in Germany and Austria .....	560
PC 164 - The Swiss Experience with Evaluation Standards .....	561
PC 165 - Evaluation Standards and Competencies: An Australasian Perspective .....	562
PC 166 - The Perspective of the UNEG .....	563
PC 167 - Competency Development and Standards in WFP .....	564
Rethinking Evaluation Approaches and Skills for Development and Human Rights Programmes in Context of Fragility, Protracted Conflict and Structural Violence .....	565
PC 169 - Challenges to Incorporating Gender When Evaluating Conflict .....	566

PC 290 - Seven Principles of Evaluating in contexts of Fragility, Conflict and Violence .....	566
PC 170 - Methodological Challenges and Evaluators Competences to Evaluate Resilience Oriented Programs. Lessons Learned from De Latinamerican Región .....	566
A Framework for Accelerating Access to Science, Technology and Evaluation Results (FAASTER) for Global Health Impact .....	567
PC 171 - Panel Contribution .....	567
PC 172 - Panel Contribution .....	568
PC 173 - Panel Contribution .....	568
Data Collection in Dynamic, Fragile and Complex Environment: Learning from the Field .....	569
PC 174 - Community Based Information Collection and Support System in Afghanistan .....	569
PC 175 - Participatory Statistics to Examine the Impact of Interventions to Eradicate Modern Slavery .....	570
Our Learnings from Ten Years' Experience Using Three Different Kinds of Evaluative Rubrics in Our Evaluation Practice: Practical Tips	571
PC 180 - An Overview of What Evaluative Rubrics Are .....	572
PC 181 - Exploring Three Different Kinds of Evaluative Rubrics .....	572
PC 182 - Reflecting on the Strengths and Weaknesses of the Different Kinds of Evaluative Rubrics .....	573
PC 183 - How We Might Use Evaluative Rubrics to Integrate Shared Values into Evaluation .....	573
Using Outcome Harvesting to Learn about what Supports Resilience .....	574
PC 184 - Lessons Learned from Piloting Outcome Harvesting for Free Press Unlimited (FPU) in Somalia and Pakistan .....	575
PC 185 - Using Outcome Harvesting to Describe Results, Support Learning and Evaluate a Civil Society Network for Disaster Reduction .....	575
PC 186 - Adapting the Outcome Harvesting tool for the systematic evaluation of multiple outcomes of agricultural research for development .....	575
Evaluation Criteria for the 21 <sup>st</sup> Century: Panel 1 Rethinking Our Evaluation Criteria .....	577
PC 187 - Criteria, Values and the Logic of Evaluation .....	578
PC 188 - Evaluation Criteria: Perspectives from the Global South .....	578
PC 189 - Criteria for the Evaluation of Sustainable Development .....	578
PC 190 - Criteria on the Interface of Private Sector and Development Evaluation .....	578
PC 191 - Discussant: Challenges for the Evaluation Profession in Redefining Criteria .....	578
Evaluation in Turbulent Times. Pushing Boundaries Through a Complexity Lens .....	579
PC 192 - Who Is Benefiting? An Approach to Measure Indirect Effects .....	580
PC 193 - Navigating Complexity in the Real Time Evaluation of Humanitarian Response to European Migration .....	580
Making Sense of Outcomes: Using the Signification Process of SenseMaker to Enrich Outcome Mapping and Outcome Harvesting	581
PC 194 - Signification of Outcomes: Enhancing The Harvesting and Analysis of Outcomes .....	582
PC 195 - Using Signifier Questions in the Analysis of the Outcome Harvesting Evaluation of Actionaid Denmark's Tax Justice Programme .....	582
Understanding How Adaptive Management Can Support Resilience Strengthening .....	583
PC 196 - Panel Contribution .....	583
PC 198 - Panel Contribution .....	584
PC 199 - Panel Contribution .....	584
Issues and Challenges in the Establishment and Development of Voluntary Organisations for Professional Evaluation (VOPE) .....	585
PC 200 - Challenges and Strategies of the Turkish Monitoring and Evaluation Society .....	586
PC 201 - Challenges and Strategies of the Jordanian Development Evaluation Association .....	586
PC 202 - Scholarship and Practical Resources for Voluntary Organisation for Professional Evaluation (VOPE) Development .....	587
PC 203 - Research and Scholarship on Organisational Evaluation Capacity Building: Insights for Developing Voluntary Organizations for Professional Evaluation (VOPEs) .....	587
EvalAgenda 2020 in the Region Eastern Europe, Central Asia & South Caucasus: Opportunities and Challenges for Young and Emerging Evaluators .....	588
PC 204 - Creating the Enabling Environment in ECA Region: the Role of YEE .....	589
PC 205 - YEEs as Drivers of Changes in ECA VOPEs .....	589
PC 206 - Development of Individual Capacities for Evaluation: Needs, Challenges, Opportunities .....	590

PC 207 - Creating Strong Interlinkages: EvalYouth ECA as the Initiative of YEE in the Region and for YEE in the Region. ....	590
PC 208 - Discussion of the Panel .....	591
Evaluation for Market Resilience: Methodological Challenges and Approaches .....	592
PC 209 - Crisis Response and Crisis Prevention .....	592
PC 210 - What Role can MDBs Play in Promoting Market Resilience? .....	593
PC 211 - Inclusion and Resilience .....	593
PC 212 - Market Resilience in Private Sector Evaluation. ....	593
The Power of Diversity and Inclusiveness (DI): A Path to Peace and Prosperity .....	594
Evaluating Unintended Effects and Adapt in Countries Affected by Fragility, Conflict and Violence .....	595
PC 289 - Evaluating Unintended Effects and Adapt in Countries Affected by Fragility, Conflict and Violence .....	596
PC 220 - The Use of Art in Reconciliation and Memory Processes .....	596
PC 221 - Evaluating Unintended Effects .....	596
PC 222 - Addressing M&E Challenges, Technical and Personal Evaluators Competences in Complex and Conflict Contexts .....	597
Psychological Approaches in Strengthening Resilience of Sustainable Development Programmes from An M&E Perspective .....	598
PC 223 - Psychological Approaches in Strengthening Resilience of Sustainable Development Programmes from an M&E Perspective .....	599
PC 224 - Conceptualising Behaviour Science Within Sustainable Development[SD] Perspective .....	600
PC 225 - Behaviour Science Principles Models and Environmental Sustainable Development .....	601
PC 226 - Behaviour Science Aspect of Mitigating Factors in Climate Change: Towards Sustainable Communities .....	601
PC 227 - Monitoring and Evaluation Aspects of Behaviour and It's Link to Sustainable Circumstances .....	602
Using Big Data to Evaluate an Experimental Program Addressing Poverty, Climate Change and Resilience among Small Farmers in SE Asia .....	604
PC 228 - The Sticky Rice Platform for Addressing Poverty, Climate Change and Resilience Among Small Farmers in Se Asia .....	605
PC 229 - Exploring Research Design Options for the Sticky Rice Program to Integrate Big Data and Conventional Evaluation Designs: Opportunities and Challenges .....	605
PC 230 - Discussant .....	606
Participatory Approaches to Resilience Measurement: Who are we Listening to? .....	607
PC 231 - Rethinking Evaluation Methods and Methodologies. ....	608
PC 232 - Participatory Approaches to Resilience Measurement: Who are we Listening to? Context Matters .....	608
PC 233 - Participatory Approaches to Resilience Measurement: Who are we Listening to? Do we Consider Gender and Other Inequalities? .....	608
PC 234 - Participatory Approaches to Resilience Measurement: Who are we Listening to? Whose Participation and How to Address Conflict of Interest? .....	609
PC 235 - Panel Contribution .....	609
Applying Theory To Policy + Program Evaluation in Different Contexts: An Integrated Approach .....	610
PC 236 - Applying Theory To Policy + Program Evaluation in Different Contexts: An Integrated Approach .....	611
PC 237 - Policy and Programme Integration in Natural Resources Focusing on the Energy Sector (Including Innovation) .....	612
PC 238 - How Theory of Change Can Work to Link Programme to Policy in Conditions of Adaptive Management. ....	612
From Niche to Mainstream? Critical Reflection on the Sensemaker Practice to Date .....	613
PC 239 - What Is the Added Value of Sensemaker? .....	614
PC 240 - What Do We Learn from Practice to Date? .....	614
PC 241 - From Niche to Mainstream? .....	614
Evaluation in Resettlement: Accountability, Vulnerability and Contested Contexts .....	615
PC 242 - Topic of Discussion: Dynamics in Resettlement and Livelihood Options for Refugees in Transition. ....	615
PC 288 - Use of Technology in Evaluations with Displaced Population in Fragility, Conflict and Violence .....	616
PC 243 - The Challenges of Monitoring and Evaluation in Resettlement in a Calamity Affected Road Project when Design is Flawed .....	616
PC 244 - Evaluating the Negotiation Space in Resettlement .....	617
National Evaluation Policies and Systems in Europe: Accomplishments and Challenges for Collective Action .....	618
PC 245 - National Evaluation Policies in Europe: Observations and Learning from the Voluntary Nations Reviews (VNRs) (Kassem) ..	618

PC 246 - Prospects for Developing a National Evaluation Policy in Greece: the Role of the Hellenic Evaluation Society .....	619
PC 287 - Towards Evidence-Based Policy Making – Institutionalization of Evaluation in Finland .....	619
<b>Navigating Spectacular Ambitions: The Importance of Thinking Evaluatively in Challenging the Sustainable Development Goals of “No One Left Behind” .....</b>	<b>620</b>
PC 247 - Incorporating Developmental Trajectories into Sustainable Development Goals .....	621
PC 248 - Questioning the Rhetoric of “Leaving No one Behind” from a Sustainability Lens. ....	621
<b>PC 249 - Some Evaluation Questions to Help Sharpen the Focus of Sustainable Development Goals in Addressing Inequities: The Different Roles for Evaluators .....</b>	<b>622</b>
Criteria for Measuring Resilience: What Do We Know and What are We Missing? .....	623
PC 250 - Refocusing the DAC Criteria Through a Resilience Lens .....	623
PC 251 - Using Development Principles to Guide Resilience Measurement .....	624
PC 252 - Resilience Measurement on the Ground: Experience From GRP’s Grantees. ....	624
<b>The use of Causal Mechanisms in Evaluation: Collecting Practical Experiences of Applying Different Mechanism-Based Approaches .....</b>	<b>625</b>
PC 253 - Using Mediation Analysis to Uncover Psychological Mechanisms of Attitude Change in a Development Volunteer Program	626
PC 254 - Experimenting with Process Tracing Under Real-World Evaluation Constraints .....	627
PC 255 - Causal Mechanisms in Theory Based Impact Evaluation. ....	627
PC 256 - Using Realist Synthesis to Document Patterns of Effectiveness for Climate Adaptation Projects in Ethiopia .....	628
Emerging Evaluation Communities in South East Europe (SEE) .....	629
PC 257 - Panel Contribution .....	630
PC 258 - Panel Contribution .....	630
PC 259 - How to Help Evaluation Communities Emerge when the Policy Stakeholders are Reluctant to Perform Evaluation? .....	630
PC 260 - Panel Contribution .....	631
Whose Needs Count? Exploring Different Priorities and Perspectives for Evaluating Resilience .....	632
PC 261 - Empowering People and Building Resilience to Risk Through the Lens of Inclusion .....	632
PC 262 - Whose resilience? The Rationale for Accounting for Systemic Inequalities, Gender in Particular, and Their Intersections in Impact Evaluations .....	633
How Can Evaluations of Governance Experiments Make Societal Transitions Manageable? .....	634
PC 263 - Phenomenal Evaluation Culture? Or Responding to Phenomenon-based Experimental Governance Through More Systemic and Responsive Evaluation .....	635
PC 264 - The Influence of Evaluations in Polycentric Settings: a Literature Review. ....	635
PC 265 - Evaluation of Service Experiments Using Human Centered Value Criteria .....	635
PC 266 - Frozen Experiments and the (Lack Of) Evaluability of Policy Instruments: A Case Study of Danish Work Place Assessments Under EU. ....	635
PC 267 - Evaluating Experiments for Polycentric Climate Governance. ....	636
90-minute Introduction to Principles-Focused Evaluation .....	638
PC 269 - Outcome Harvesting – A Principles-Focused Evaluation Approach. ....	639
Visualising Theories of Change .....	640
PC 270 - Representing Theories of Change: A Technical Challenge with Evaluation Consequences. ....	640
PC 271 - To the Promised Land: The case of The Salvation Army on its Quest for the ‘Perfect’ Theory of Change .....	641
PC 272 - Counting What Counts: Value in Theories of Change .....	641
PC 273 - More Ideas for Making Theories of Change More Informative (And Robust) Using Data Visualization .....	642
United Nations’ Evaluation Functions for More Resilient Societies .....	643
PC 274 - Panel Contribution .....	644
PC 275 - Panel Contribution .....	644
PC 276 - Panel Contribution .....	644
PC 277 - Panel Contribution .....	644
PC 278 - Panel Contribution .....	644
PC 279 - Panel Contribution .....	644

Lessons Learnt in Including Young and Emerging Evaluators for a Resilient Field. A Cross Regional Exchange .....	645
PC 291 - The Experience of Including Young and Emerging Evaluators from Evalyouth Costa Rica .....	646
PC 280 - Perspective from Africa. Experiences, Successes, Challenges and Motivations for Continued Evaluation Capacity Building Amongst Young and Emerging Evaluators .....	647
PC 281 - Creating Opportunities for Emerging Evaluators for a Better Evaluation Field. The Ideas Approach .....	647
PC 282 - Presenting the EES Thematic Working Group for Emerging Evaluators (TWG5). The Horizon for Emerging Evaluators in Europe	648
From Independent Evaluator to Learning Partner: Reflections on the Journey .....	649
PC 283 - Panel Contribution .....	649
PC 284 - Panel Contribution .....	649

## Round-tables

650

Focusing Sessions: A Stakeholder Engagement and Evaluation Design Method .....	651
RC 01 - Round-table Contribution .....	651
RC 02 - Round-table Contribution .....	651
When Focusing on Results Makes Impact Less Likely, and When it Doesn't .....	652
RC 03 - When Focusing on Results Makes Impact Less Likely .....	652
RC 04 - When Focusing on Results Could Make Impact More Likely .....	653
RC 05 - Moderator .....	653
Evaluations in Fragile Contexts: How Do They Differ? .....	654
RC 06 - Evaluations in Fragile Contexts: How Do They Differ? .....	655
RC 07 - Evaluations in Fragile Contexts: How Do They Differ? .....	655
RC 08 - Evaluations in Fragile Contexts: How Do They Differ? .....	655
RC 09 - Evaluations in Fragile Contexts: How Do They Differ? .....	655
Developing Synergies: the Evaluation Community and the Housing Sector .....	656
RC 10 - Round-table Contribution .....	657
RC 11 - Round-table Contribution .....	657
RC 12 - Round-table Contribution .....	657
RC 13 - Round-table Contribution .....	657
RC 14 - Round-table Contribution .....	657
RC 15 - Round-table Contribution .....	657
Digital – a Need to Have, not a Nice to Have in Evaluation. ....	658
RC 16 - Round-table Contribution .....	658
RC 17 - Round-table Contribution .....	659
RC 18 - Round-table Contribution .....	659
Institutionalization of Evaluation in Europe: From Boom Times to New Turbulences? .....	660
RC 19 - Round-table Contribution .....	661
RC 20 - Round-table Contribution .....	661
RC 21 - Round-table Contribution .....	662
RC 22 - Round-table Contribution .....	662
How Is Resilience Measurement Moving us Beyond Just Good M&E?: The State of Resilience Evaluation Practice and the Paths Ahead .....	663
RC 24 - Evaluating Large-Scale Climate Resilience Investments: What Does Complexity Theory Bring? .....	664
RC 25 - Resilience Measurement: Major Achievements and Outstanding Challenges .....	664
RC 26 - Evaluating Resilience in Urban Areas and Populations .....	664
RC 27 - Round-table Contribution .....	664

Joint Session of the TWG on Private Sector Evaluation (PSE) and the AEA TIG on Social Impact Measurement (SIM) . . . . .	665
RC 30 - Thematic Working Group on Private Sector Evaluation. . . . .	666
RC 31 - Evaluation of PPPs. . . . .	666
RC 32 - Presentation of the Topical Interest Group(TIG) on Social Impact Measurement of the American Evaluation Association (AEA) . . . . .	666
RC 33 - Utilizing Bottom-Up Approach in Social Impact: Going from Community Entrepreneur to Impact Investor . . . . .	666
RC 34 - Messages from the Social Impact Measurement Arena . . . . .	667
The Road Ahead for Evaluation in the Lusophone World: Resilience and Action in Critical Times . . . . .	668
RC 35 - AfDB's Evaluations in Lusophone countries in Africa: Experience, Challenges and Opportunities on the Way Forward. . . . .	669
RC 36 - What Can We Learn from the Institutionalization of M&E in Brazil?. . . . .	669
RC 37 - Building Evaluation Capacity: Early Lessons from the WACIE Initiative. . . . .	669
RC 38 - Present and Future Challenges of the Evaluation in Portugal. . . . .	670
RC 39 - The Policy-Maker Perspective: What Are the Bottlenecks to Evaluation in Cabo Verde and Africa . . . . .	670
Scrambling after Moving Targets: M&E Applied to Adaptive Management in Peacebuilding . . . . .	671
RC 40 - Connect. Learn. Adapt. Repeat. Using Adaptive Management to Bolster Evaluation and Resiliency in Conflict-Affected Areas . . . . .	672
RC 41 - Existing Evaluation Approaches and How They Fit with Adaptive Management . . . . .	672
At the Axiological Intersection of Epistemology and Ontology: Or, Where the Rubber Hits the Road. . . . .	674
RC 42 - At the Axiological Intersection of Epistemology and Ontology: Or, Where the Rubber Hits the Road. The Former Student/Current Instructor Perspective . . . . .	675
RC 43 - At the Axiological Intersection of Epistemology and Ontology: Or, Where the Rubber Hits the Road. The Original Program Instructor Perspective . . . . .	676
RC 44 - At the Axiological intersection of Epistemology and Ontology: Or, Where the Rubber Hits the Road – The Management Perspective. . . . .	677
RC 97 - At the Axiological Intersection of Epistemology and Ontology: Or, Where the Rubber Hits the Road. The Program Supervisor Perspective . . . . .	678
What Lasts? Exploring Sustainability: What Does It Mean? And How Can It Be Measured? . . . . .	679
RC 45 - Critical Times and Critical Questions: What Does Sustainability Mean for Climate Change Focused Interventions . . . . .	680
RC 46 - Evaluating Sustainability in Relation to Children's Rights. . . . .	680
RC 47 - Children Rights and Gender: Where the Two Roads Meet . . . . .	680
RC 48 - Solutions and Sustainable Reintegration: What Are We After and How Do We Identify It?. . . . .	681
RC 49 - Sustainability of Gender Equality Interventions . . . . .	681
Accountability in Evaluation for Sustained and Resilient Impacts. . . . .	682
RC 50 - Accountability in Evaluation for Sustained and Resilient Impacts: the Funders' View . . . . .	683
RC 51 - Accountability in Evaluation for Sustained and Resilient Impacts: A View from NGOs and Participants. . . . .	683
RC 52 - Round-table Contribution . . . . .	683
New Frontiers in Development Evaluation: The Case of Blended Finance. . . . .	684
RC 53 - The OECD DAC Blended Finance Principles for Unlocking Commercial Finance for the SDGs . . . . .	685
RC 54 - Methodological Issues on the Evaluation of Blended Finance . . . . .	685
RC 55 - Evaluating the Use of Subsidies in a Development Bank . . . . .	685
RC 56 - Independent Evaluation of EU Blending Facilities (2007–2013) . . . . .	685
RC 57 - Insights and Lessons Learned from Evaluations of Impact Investment Funds . . . . .	686
Meaningful Monitoring and Evaluation Systems Through Evidence-Based Leadership and Evolutionary Purpose . . . . .	687
RC 58 - Part 1: How to Build a Career in the Evaluation Field . . . . .	688
RC 59 - Part 2: Self- Assessment of Evaluators Competences . . . . .	688
RC 60 - Part 3: Formulation of a Personal Development Plan. . . . .	688
RC 61 - Part 4: Useful resources . . . . .	689
Evaluation Capacity Development: a Reflection on the Past, Present and Future of the International Program for Development Evaluation Training (IPDET) . . . . .	690
RC 62 - Round-table Contribution . . . . .	691

RC 63 - Round-table Contribution .....	691
RC 64 - Round-table Contribution .....	691
How to Build a Career in Evaluation?: Assess Your Competencies and Develop a Professional Development Plan. Roundtable or Unconference Event .....	692
RC 65 - How to Build a Career in Evaluation?: Assess Your Competencies and Develop a Professional Development Plan. Roundtable or Unconference Event .....	693
RC 66 - How to Build a Career in Evaluation?: Assess Your Competencies and Develop a Professional Development Plan. Roundtable or Unconference Event .....	694
The System for Evaluating International Development Cooperation in Sweden: Is the System Fit for Purpose? .....	696
RC 67 - The System for Evaluating International Development Cooperation in Sweden: Is the System Fit for Purpose? .....	697
RC 68 - The Perspective of the Swedish International Development Cooperation Agency (SIDA) .....	697
RC 69 - The Perspective of the Expert Group for Aid Studies (EBA) .....	698
RC 70 - The Perspective of the Swedish Ministry of Foreign Affairs .....	698
RC 71 - An Outside Perspective from DEval in Germany .....	698
RC 72 - The Perspective of FCG Sweden, an Evaluation Provider in the System .....	698
How to Demonstrate Value for Money? Methods and Trends .....	699
RC 73 - The Importance of Engaging Stakeholders in Social Outcomes Monetization .....	700
RC 74 - Economic Evaluation .....	700
RC 75 - Value for Money Approaches and The Private Sector .....	700
RC 76 - Impact Measurement in the Private sector .....	701
RC 78 - Value for Money in the Field of Democracy and Rule of Law .....	701
RC 79 - Value for Money in the Context of INGO's .....	701
RC 80 - The Value for Money agenda: A missed opportunity .....	701
Contribution of Meta-Studies for Building Resilient Societies .....	703
RC 82 - Round-table Contribution .....	704
RC 83 - Round-table Contribution .....	704
RC 84 - Round-table Contribution .....	704

## Posters

705

P 01 - Capacity to Do and to Use Evaluation in Community-Based Organizations in Québec (Canada): Development and Validation of a Measurement Instrument .....	706
P 02 - Challenges in Evaluating Small-scale Projects with Limited Resources .....	707
P 03 - In Which Way Does Evaluation Help to Make Government More Open and Accountable? .....	708
P 04 - Strategic Analysis of the Mais Médicos Program: A Performance Assessment .....	709
P 05 - Advanced Evaluation Tools and Methods for Developmental Interventions: Case Study of Sustainable Impact of a Scholarship Programme on Recipients .....	710
P 06 - Evidence on Adolescent Girls' Aspirations in South Sudanese Refugee Camps in Ethiopia .....	711
P 07 - Potentials and Challenges in the Analysis of Health Governance Policy Networks in Two Regions of the Legal Amazon - Brazil .....	712
P 08 - Determinants to Care-Seeking at Public Health Facilities in Rural Bihar and Uttar Pradesh .....	713
P 09 - Impact Evaluation in Global Innovation Hubs .....	714
P 10 - Impact of Multinational NERICA Rice Dissemination Project on Socio-economic Life of Rice Farmers in Ekiti State, Nigeria .....	715
P 11 - Lost Road: A Dilemma between Independence and Participation of Evaluation in South Korea .....	716
P 12 - Effect Decomposition in Prevention Research: Modern Causal Inference Methods .....	717
P 13 - Practical Data Collection Tools for Engaging Stakeholders in Impact Measurement .....	718
P 14 - Promoting Resilience In Partnerships Through The Use of an Appreciative Inquiry Evaluation .....	719
P 15 - The Hopscotch of Evaluation, an Integrative Methodological Approach .....	720

P 16 - “AcolheSUS Project” at the Dirceu Arcoverde Hospital: use of Co-Management Devices to Promote Changes in Institutional Reality .....	721
P 17 - Innovative Methods for Useful Results, by the Center of Excellence for Development, Impact and Learning (CEDIL) .....	722
P 18 - Finding the Balance of Power Dynamics Between Donors & Ngos, a Perspective of an Emerging Evaluator in the Global South	723
P 19 - The Perils of Selecting and Implementing Evidence-Based Programs in Behavioral Health Sciences .....	724
P 20 - A Review of Scandinavian Donors Evaluation Reports in African Interventions: Methodological Lessons .....	725
P 21 - Qualitative Approaches for Evaluating a Third-Party Payment Mechanism for Obstetric Care in the Culturally Sensitive Context of Northern Cameroon .....	726
P 22 - Evaluating Resilience Using Econometric Methods .....	727
P 23 - Conceptualising Behaviour Science Within Sustainable Perspective .....	729
P 24 - Patient-Provider Communication Related to Appropriate Drug Administration: Key to Improving Treatment Adherence .....	730
P 25 - Piloting of Using Quick Respond (QR) Code and Imaging Technology to Support Routine Field Monitoring and Evaluation Activities .....	731
P 26 - Impacting Social Norms Around Early Marriage Through Mass and Community Media: Lessons From Rural India .....	732
P 27 - Mixed Methods to Evaluate of the Effectiveness of Task Sharing to Community Health Workers – Evaluating: Preparedness, Safety and Effectiveness .....	733
P 28 - The Periodic Table of Evaluation .....	735
P 29 - From the Tree to the Forest: Reviewing Alkin’s Approach .....	736
P 30 - Guided Technical Visit: Bridge Device Between Multiprofessional Networks and Users, for the Mobilization of Knowledge in an Institutional Learning .....	737
P 32 - Conceptualising Behaviour Science Within Sustainable Perspective .....	738
P 33 - Evaluation of Monitoring and Evaluation Processes Institutionalization Related to National Technical Cooperation of the Ministry of Health with Federal Institutions .....	739
P 34 - A Multilayer Mechanism to Maximize Results and Embed Evaluations .....	740
P 35 - Evaluation of Good Projects: Looking for Opportunities and “Power Results” .....	741
P 36 - Measuring Effectiveness in the Experiment of Inclusive Social Security .....	742
P 37 - Evaluation Challenges in the Health Field. Lessons Learnt from Evaluating the Cervical Cancer Prevention Program in San Juan, Argentina .....	743
P 38 - Developing Partnerships for Communicating Evaluation Findings Using the Site Web: Experiences of Evidence for Development Initiative in Francophone Africa .....	744
P 39 - Communicating and Using Evaluation Results in Fragile and Conflict-Affected Countries .....	745
P 40 - Digital Platform for Monitoring and Evaluation AcolheSUS Project: A Proposal for Reflection on the Daily Practices of Health Services .....	746
P 42 - Emerging VOPE Led by Young Evaluators .....	747
P 43 - Higher Education in Forced Migration: A Comparative Analysis of Scholarship Programmes for Syrian Refugees Implemented by Giz & Daad .....	748
P 44 - Resilient Schools Embrace Social Innovation: The Case of Family Group Conferences .....	750

## Authors Index

751

# Keynote Lectures

Wednesday, 3 October 2018  
09:00–10:00

## KL 01 Keynote Lecture 1

# K 1 - Resilient Societies and Resilient Evaluators in Turbulent Times

L.E. Pleger<sup>1</sup>, F. Leeuw<sup>2</sup>

<sup>1</sup> *Zhaw School of Management and Law, Switzerland*

<sup>2</sup> *University of Maastricht, Law, Public Policy and Social Science Research, Professor, the Netherlands*

The presentation will be a collaborative speech, held by Prof. Dr. Frans L. Leeuw and Dr. Lyn E. Pleger together. It first tackles the topic of evaluations in the context of resilience from both the conceptual and empirical perspectives, by combining theories and findings from different research fields, which build upon one another. Next, it will address the relationship between resilient societies and resilient evaluators: how can and will evaluators, in turbulent and stressful times, with fakenews, evidence mazes, disbelieving evidence, global challenges and risks attributed to data-ification and digitization, be resilient and effective in their work and communication? The presentation starts by clarifying the overall theoretical framework consisting of three levels of resilience, namely [1] individual, [2] organization and [3] society at large. An overview is given of what main findings and approaches are with regards to resilience at these three levels. Resilience is then linked to evaluations in general and the need of evidence-based information in particular. In this regard, the presentation will discuss two challenges in the context of evidence-based evaluations in more detail: Firstly, challenges to produce evaluation results that 'matter'. Secondly, how information can reach the target groups of this information. Building on this, the presentation addresses the importance of evaluations from a descriptive perspective on the one hand (growth in evaluation numbers) and a content related perspective (use of evaluations) on the other. In a next step, the topics of evaluator resilience and independence of evaluations as requirements for evidence-based information are approached. These issues are considered by presenting survey findings regarding independence of evaluations and this is complemented by using behavioral and psychological findings.

Thursday, 4 October 2018  
10:15–11:00

## KL 02 Keynote Lecture 2

### K 2 - Post-normal Evaluation?

*T. Schwandt*<sup>1</sup>

<sup>1</sup> University of Illinois, Emeritus Professor of Educational Psychology, USA

Evaluation evolves, responding to developments in the technical, social, and political contexts in which it unfolds as well as influencing how those contexts are to be understood and appraised. Evaluation has typically been understood as a reliable means of providing some certainty in the appraisal of value, but increasingly has been seen as having to address a variety of uncertainties and risks in such appraisals. Some new ways of thinking and forms of evaluation practice linked to planning and decision making are emerging reflecting assumptions of unpredictability as well as incompleteness, instability, and a plurality of perspectives in value determination. This may signal the emergence of “post-normal evaluation.” This talk explores several key features of this development.

Thursday, 4 October 2018  
14:00–14:45

### KL 03 Keynote Lecture 3

## K 3 - Caroline Heider's Keynote Lecture

C. Heider<sup>1</sup>

<sup>1</sup> World Bank Independent Evaluation Group, former Director-General, USA

Caroline Heider's keynote speech aims to set the scene for the conference. It focuses on the role of evaluation in developing capacities for resilient societies. In her opening remarks, Caroline unpacks several inter-related dimensions of resilience. She refreshes the audiences' understanding of what it takes to develop capacities. And, she reflects on the duality the current context presents to evaluation. On the one hand, it provides opportunities for evaluation to support the process of building greater societal resilience. At the same time, evaluation needs to adapt to influence change and fulfill its potential.

Resilience is reviewed from a process point of view, a cycle of response and rebuilding after shocks. This process has intrinsic evaluative elements, or rather, if not undertaken evaluatively the process might lead to Sisyphus rebuilding endlessly. Three dimensions are discussed: economic, environment, and social, which are deeply intertwined and have complex multiplier effects.

Caroline builds on applied research to capacity development, which goes well beyond "training" (even if the terms are often used interchangeably) and requires a functioning system of enabling environment, institutions, and individuals. Likewise, the process of "building" capacity calls for a dynamic process, which involves analysis and evaluation, planning and serendipity.

Together, these two conceptual frameworks provide a complex web of processes and desired outcomes. They require some unpacking, while retaining connectivity, to create opportunities for evaluation. Because both – resilience and capacity development – have intrinsic evaluative elements, the evaluation practice should be well placed to support this intertwined process.

But, challenges remain to keep evaluation practices relevant and updated. Some might see these as shortcomings, others – including Caroline – revel in the opportunity to see evaluation grow as practice and profession.

Friday, 5 October 2018  
10:15–11:00

#### KL 04 Keynote Lecture 4

### K 4 - Learning About Evaluation from a Conference?

E. Stern<sup>1,2,3</sup>

<sup>1</sup> Lancaster University, Evaluation Research, Emeritus Professor, Lancashire, United Kingdom

<sup>2</sup> University of Bristol, Honorary Fellow, Bristol, United Kingdom

<sup>3</sup> *International Journal of Theory, Research and Practice*, Editor of Evaluation

Elliot Stern has been invited in a keynote towards the end of the conference to offer some synthesis and reflection on what he has been learned whilst in Thessaloniki. He plans to treat the conference as a microcosm embedded in a wider evaluation community and system. The raw material for this talk is the content of the conference, including what does or does not get covered; who is present – but also who is absent; and how this conference compares with other EES conferences as far as he remembers. The hope is that this 'data' will tell us something about contemporary evaluation agendas and priorities; the European and to an extent the global evaluation community; and the EES relationship to these communities. Elliot expects to be informed by inputs from other colleagues over the course of the conference; and to share the synthesis and reflection task with those present during his keynote.

# Flagship Symposia

Wednesday, 3 October 2018  
10:00–11:30

---

**FGS 01 FLAGSHIP SYMPOSIUM 1: Better Evaluation for better policies in Greece beyond the crisis**

## **Better Evaluation for Better Policies in Greece Beyond the Crisis**

B. de Laat<sup>1</sup>

<sup>1</sup> *European Evaluation Society, Board, Paris, France*

Wednesday, 3 October 2018  
10:00–11:30

**FGS 01 FLAGSHIP SYMPOSIUM 1: Better Evaluation for better policies in Greece beyond the crisis**

## **FS 01 - Contribution of Ms. Gerovasili**

*O. Gerovasili*<sup>1</sup>

<sup>1</sup> Minister of Citizens' Protection, Greece

**FGS 01 FLAGSHIP SYMPOSIUM 1: Better Evaluation for better policies in Greece beyond the crisis**

## **FS 02 - Contribution of Ms Louka Katseli**

*L. Katseli*

**FGS 01 FLAGSHIP SYMPOSIUM 1: Better Evaluation for better policies in Greece beyond the crisis**

## **FS 03 - Contribution of Ms Anastasia Tzavaras Catsambas**

*A. Tzavaras Catsambas*<sup>1</sup>

<sup>1</sup> EnCompass, Evaluation, Rockville, USA

Wednesday, 3 October 2018  
14:30–16:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**FGS 02 FLAGSHIP SYMPOSIUM 2: Evaluation Unusual: How is Resilience Challenging the Evaluation Field, and Are We Rising Effectively to the Challenge?**

## Evaluation Unusual: How is Resilience Challenging the Evaluation Field, and Are We Rising Effectively to the Challenge?

*K. Jon*<sup>1</sup>

<sup>1</sup> Mercy Corps, Research and Learning, Washington DC, USA

Purpose of the Symposium A high-profile session on Day 1 to set the conference alight; designed to:

- challenge and provoke conventional evaluation thinking;
- explore why and in what specific ways the resilience concept and its application requires a rethink (or not) of taken-for granted assumptions about evaluation;
- inspire the conference participants, in part by profiling major development and innovations in resilience evaluation and outlining persistent challenges.

Abstract Evaluation as usual is being challenged by the concept of resilience. Should resilience be measured as a set of adaptive capacities or as end-state? Are there common principles for evaluating resilience investments across different systems, scales, and types of shocks? Do efforts to quantify resilience run counter to understanding the complex dynamics inherent to the concept resilience? Can contributions towards resilience be evaluated in the absence of a shock? This symposium will tackle these and other evaluative questions that have surrounded the concept of resilience as it has gained meaning and power across disciplines, sectors and geographies – with a focus on its application in the humanitarian and international development fields. The dynamic discussion will elucidate how the art and science of evaluation is evolving – and needs to further evolve – to translate the concept of resilience into improve policy and practice. Format of the Symposium 90 minutes, dynamic moderator, with 4–5 evaluation specialists coming from different disciplines/ sectors that have been applying the concept of resilience. These are evaluation specialists who provide leadership in innovations and partnerships that have come about in order to address the evaluation challenges that the resilience concept poses. The format will be a robust conversation (round table format) informed by 4–5 critical questions that are fundamental to this field. For example:

- What are the unique characteristics of the resilience concept/ resilience theory that challenge the art/ science of evaluation?
- How are evaluators managing the inherent complexity in measuring and evaluating resilience to generate actionable insights and transferable evidence?
- What has the experience of resilience measurement at different scales (individual, household, community, systems) taught us about cross-scale measurement within resilience and in other fields?
- Does getting down to evaluating the component parts of systems and capacities within real-time frames reduce resilience measurement to being an emperor with no clothes – is this just evaluation as usual?
- What is the emerging evidence from evaluations of resilience initiatives telling us? Where are the gaps, and how can evaluation professionals address these?

Wednesday, 3 October 2018  
14:30–16:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**FGS 02 FLAGSHIP SYMPOSIUM 2: Evaluation Unusual: How is Resilience Challenging the Evaluation Field, and Are We Rising Effectively to the Challenge?**

**FS 04 - Contribution of Tim Frankenberger**

*T. Frankenberger*<sup>1</sup>

<sup>1</sup> TANGO International, President, Tuscon, USA

The 4 speakers are experienced evaluators, who are developing methods and tools in response to the challenges being faced in different disciplines and sectors, in the application of the concept of resilience to inform decision-making by planners and investors.

This symposium offers an opportunity to identify the differences between these disciplines/ sectors when it comes to the ways in which evaluation as-usual is being challenged by taking on the concept of resilience, and an exciting moment in which clear synergies and bridges can be surfaced that can be of benefit (e.g. in what ways can urban resilience evaluation learn from resilience evaluation approaches applied to recurrent food security crises in drought emergencies?), and common principles can be identified.

The speakers will include: Tim Frankenberger, TANGO International. Will draw on experience in address the challenges of evaluating resilience as developed and applied in the context of food/ livelihood security and 'ending drought emergencies' (Horn of Africa, Sahel)

**FGS 02 FLAGSHIP SYMPOSIUM 2: Evaluation Unusual: How is Resilience Challenging the Evaluation Field, and Are We Rising Effectively to the Challenge?**

**FS 05 - Contribution of Jyotsna Puri**

*J. Puri*<sup>1</sup>

<sup>1</sup> Green Climate Fund, Independent Evaluation Unit, Yeonsu-gu-Incheon, Republic of Korea

Jyotsna Puri, Green Climate Fund. Will draw on experience of working on complexity (particularly related to humanitarian crises) and evaluation of resilience.

**FGS 02 FLAGSHIP SYMPOSIUM 2: Evaluation Unusual: How is Resilience Challenging the Evaluation Field, and Are We Rising Effectively to the Challenge?**

**FS 06 - Contribution of Maliha Khan**

*M. Khan*<sup>1</sup>

<sup>1</sup> Daira, Independent Consultant, London, United Kingdom

Maliha Khan, Daira. Will talk draw on experience of developing measurement/ evaluation systems to capture resilience dynamics between and across scales, as particularly arising in the context of urban resilience.

**FGS 02 FLAGSHIP SYMPOSIUM 2: Evaluation Unusual: How is Resilience Challenging the Evaluation Field, and Are We Rising Effectively to the Challenge?**

**FS 07 - Contribution of Claire Hutchings**

*C. Hutchings*<sup>1</sup>

<sup>1</sup> Oxfam GB, Head of Programme Quality, Oxford, United Kingdom

Claire Hutchings or Irene Guijt. Will speak to approaches to analyzing complex systems changes, drawing on evaluation and sensemaking approaches to resilience programs and advocacy campaigns.

Thursday, 4 October 2018  
17:30–19:00

**FGS 03 FLAGSHIP SYMPOSIUM 3: Evaluation Criteria for the 21<sup>st</sup> Century: Panel 2 Revisiting the DAC Criteria**

## Evaluation Criteria for the 21<sup>st</sup> Century: Panel 2 Revisiting the DAC Criteria

*Z. Ofir<sup>1</sup>*

<sup>1</sup> Stellenbosch University, Independent Evaluation Specialist, Tygerberg, South Africa

This panel will follow on from Panel 1. It will focus on the rationale for, and emerging perspectives from the DAC criteria review launched by the DAC Network on Development Evaluation in March 2018. The panel members will discuss what has been learnt from the application of the criteria in conventional, development and humanitarian evaluation, and seek views from conference participants on the following issues:

- (i) Is a 'generic' (standard) set of criteria in common use essential or even desirable?
- (ii) What are the main lessons from the practical use of the DAC criteria over the past fifteen years?
- (iii) In the existing set of criteria, what needs updating or refining?
- (iv) Where are gaps and what are emerging suggestions for new criteria?
- (v) How can we ensure that a new or revised set of criteria is well implemented and does not fall into the trap of mechanistic application?

The floor will then be open for comment and debate in the light of the perspectives presented by both panels, including with the purpose of getting further input that can inform the DAC criteria revision process.

Thursday, 4 October 2018  
17:30–19:00

**FGS 03 FLAGSHIP SYMPOSIUM 3: Evaluation Criteria for the 21<sup>st</sup> Century: Panel 2 Revisiting the DAC Criteria**

**FS 08 - N/A - Interactive Discussion**

H.E. Lundgren<sup>1</sup>

<sup>1</sup> OECD DAC Evaluation Network, Manager, Paris, France

Hans Lundgren's contribution How each panelist will contribute and for how long: An initial 40 min anchoring interactive discussion between the Chair and panel members will be followed by an open discussion, including with the purpose of soliciting further input into the DAC criteria review process.

**FGS 03 FLAGSHIP SYMPOSIUM 3: Evaluation Criteria for the 21<sup>st</sup> Century: Panel 2 Revisiting the DAC Criteria**

**FS 09 - N/A - Interactive Discussion**

C. Heider<sup>1</sup>

<sup>1</sup> IEG- World Bank Group, Director General and Senior Vice President- Evaluation, Washington, USA

**FGS 03 FLAGSHIP SYMPOSIUM 3: Evaluation Criteria for the 21<sup>st</sup> Century: Panel 2 Revisiting the DAC Criteria**

**FS 10 - N/A - Interactive Discussion**

P.O. Bastoe<sup>1</sup>

<sup>1</sup> Norwegian Agency for Development Cooperation Norad, Evaluation Director, Oslo, Norway

**FGS 03 FLAGSHIP SYMPOSIUM 3: Evaluation Criteria for the 21<sup>st</sup> Century: Panel 2 Revisiting the DAC Criteria**

**FS 11 - N/A - Interactive Discussion**

T. Schwandt<sup>1</sup>

<sup>1</sup> University of Illinois, Educational Psychology, Urbana a Champaign, USA

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**FGS 04 FLAGSHIP SYMPOSIUM 4: The Work of the European Commission Regulatory Scrutiny Board - Building Institutional Resilience**

B. Naudts<sup>1</sup>

<sup>1</sup> *European Commission, Regulatory Scrutiny Board, Brussels, Belgium*

The Commission's 2015 Better Regulation agenda aims at delivering well-targeted, evidence-based regulation that is more likely to be properly implemented and achieve its goals. Ex-ante and ex-post evaluation are essential tools in the Better Regulation approach. Ensuring that legislation remains relevant, future-proof and effective, at the lowest possible cost should produce policies that are more resilient to societal change.

In this framework, the Regulatory Scrutiny Board (RSB) was created to improve the quality of evaluations and impact assessments. As such, it is one of the key instruments to institutionalise evidence-based policymaking and to promote institutional resilience.

The Flagship symposium will explain the role and functioning of the RSB. It will try to assess its influence on the quality of evaluations and impact assessments, and on the content of policy proposals. A number of evaluation and impact assessment examples in the domains of environment, transport, taxation and humanitarian aid will be presented to illustrate how evaluation and scrutiny work in practice in the European Commission.

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**FGS 04 FLAGSHIP SYMPOSIUM 4: The Work of the European Commission Regulatory Scrutiny Board - Building Institutional Resilience**

## **FS 12 – Symposium Contribution**

V. Valant<sup>1</sup>

<sup>1</sup> Slovenia

**FGS 04 FLAGSHIP SYMPOSIUM 4: The Work of the European Commission Regulatory Scrutiny Board - Building Institutional Resilience**

## **FS 13 – Symposium Contribution**

G. Ebling<sup>1</sup>

<sup>1</sup> European Commission, Brussels, Belgium

**FGS 04 FLAGSHIP SYMPOSIUM 4: The Work of the European Commission Regulatory Scrutiny Board - Building Institutional Resilience**

## **FS 14 – Symposium Contribution**

J. Nilsson<sup>1</sup>

<sup>1</sup> Sweden

**FGS 04 FLAGSHIP SYMPOSIUM 4: The Work of the European Commission Regulatory Scrutiny Board - Building Institutional Resilience**

## **FS 15 - Symposium Contribution**

M. White

# Fishbowl Sessions

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**FBS 01 Promoting Gender Responsive and Equity Focused Evaluation among Young and Emerging Evaluators: Experiences in Cross Collaboration between EvalYouth and EvalGender+**

## Promoting Gender Responsive and Equity Focused Evaluation among Young and Emerging Evaluators: Experiences in Cross Collaboration between EvalYouth and EvalGender+

C. Olavarria<sup>1</sup>, M. Segone<sup>2</sup>, G. Sanchez<sup>3</sup>, S. Negroustoueva<sup>4</sup>, R.B. Nanda<sup>5</sup><sup>1</sup> EvalYouth LAC, Research, Santiago, Chile<sup>2</sup> United Nations Population Fund, Evaluation Office, New York City, USA<sup>3</sup> EvalYouth LAC, Research, Mexico City, Mexico<sup>4</sup> EvalGender+, Co chair, Washington DC, USA<sup>5</sup> EvalGender+, Community of Practice, New Delhi, India

EvalGender+ and EvalYouth are two networks that operate under EvalPartners partnership to foster knowledge sharing and networking among M&E practitioners worldwide. In December 2017 EvalYouth and EvalGender+ representatives, from Africa, Asia and America gathered in México to develop a formal strategy to enhance a more effective implementation of collaboration between both networks. Many possibilities have been identified: engaging of and strengthening skills of young and emerging evaluators and promoting the demand and use of EFGRE. Three objectives were prioritized: (1) capacity development for inclusion of a gender perspective in evaluation among young and emerging evaluators, (2) increasing the sustainability of the collaboration, and (3) mentoring experience for young evaluators in an EFGRE collaboration. In this round table, participants will reflect on key lessons from operationalizing these objectives to-date, review the possibilities that emerge from the context-specific experience, and draw paths for further and deeper collaboration with and through EES members. The Global Evaluation Agenda 2020 shows the priorities and needs to strengthen the evaluation field at the global level. However these challenges have their correlation in the regional and local levels, with other specific contextual challenges. In this frame, discussing the lessons learned from the joint work of different networks at the global level opens a wide range of possibilities to explore cross collaboration between other EvalPartners initiatives, thematic groups, VOPEs and/or other stakeholders of the evaluation field at different territories and levels. Allowing an open space to discuss the lessons learned on the implementation of the three planned initiatives between EvalGender+ and EvalYouth- joint webinar, joint taskforce and a hands on mentoring experience- opens the debate and discussion about new innovative ways to collaborate and achieve the goals of the EvalAgenda 2020 and find new ways for further and deeper collaboration in the global, the regional and the national level

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**FBS 01 Promoting Gender Responsive and Equity Focused Evaluation among Young and Emerging Evaluators: Experiences in Cross Collaboration between EvalYouth and EvalGender+**

## **FB 01 - It is Never Too Early or Too Late to Learn About EFGRE: Co-Chairs Experiences with Capacity Development**

*F. Amariles<sup>1</sup>*

<sup>1</sup> *Learning for Impact, Chief Executive Officer, Weston, USA*

Svetlana will share learning from a compilation of hands-on capacity building efforts to introduce and mainstream equity-focus and gender-responsive M&E in primarily seasoned professionals in their respective fields, such as climate change, health and even feminist advocacy, across a range of funders. The session will focus on experiences from teaching how to identify entry points for EFGRE opportunities in M&E, challenges and successes.

**FBS 01 Promoting Gender Responsive and Equity Focused Evaluation among Young and Emerging Evaluators: Experiences in Cross Collaboration between EvalYouth and EvalGender+**

## **FB 02 - Contribution of Rituu B. Nanda**

*R.B. Nanda<sup>1</sup>*

<sup>1</sup> *Eval Gender+, Community of Practice, New Delhi, India*

Through this presentation, Rituu will share her experience of engaging youth in evaluations. The presentation will shed light on the factors which helped in engaging the youth and the challenges faced in the process. Rituu will expand on what was the result of engaging youth particularly from marginalized communities in evaluation. The presenter would like to use this opportunity to seek inputs from the attendees on ways to get the voice of the communities in evaluation so that 'no one is left behind' and to ensure that this is a meaningful process and not tokenistic or limited to consultation

**FBS 01 Promoting Gender Responsive and Equity Focused Evaluation among Young and Emerging Evaluators: Experiences in Cross Collaboration between EvalYouth and EvalGender+**

## **FB 03 - A Local Perspective of Evalgender and Evalyouth Collaboration; The Latin American Experience**

*G. Sánchez Romero<sup>1</sup>*

<sup>1</sup> *EvalYouth LAC, Research, Mexico City, Mexico*

2015 was declared as the international year of Evaluation by EvalPartners. Evaluation delegate from the five continents created the Evaluation Agenda 2016–2020. This Agenda stated that evaluation has to be conducted with the inclusive vision of "No one left behind". In the light of this principle, EvalPartners launched, amongst others, two initiatives addressed at including and strengthening young and emerging evaluators around the globe called "EvalYouth", while the other was aimed at including a gender perspective and gender balance into the practice of evaluation. The regional chapters have since been successful at conducting activities in line with these initiatives locally. In 2017, at the International Evaluation for the SDGs conference in Guanajuato, Mexico, a working group was organized between EvalGender and EvalYouth, with the goal of drafting an agenda that promotes the inclusion of both young and female

evaluators in teams. In consequence, national networks of evaluators around the world have gained interest in incorporating these perspectives in their on practices and some have conducted activities aligned with these initiatives. Thus, this paper will present different examples of ho national networks of evaluation have put into practice this agenda in Latin America. We will also discuss the main achievements and best practices, but also the challenges and opportunities to make this agenda happen.

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**FBS 03 EVALSDGs - Evaluation Is an Essential Ingredient for the SDGs - Sharing Experiences**

## **EVALSDGs - Evaluation is an Essential Ingredient for the SDGs – Sharing Experiences**

*D. Lucks<sup>1</sup>, A. Ocampo<sup>2</sup>, S. DErrico<sup>3</sup>*<sup>1</sup> *SDF Global, EVALSDGs, Barragup, Australia*<sup>2</sup> *UNICEF, Evaluation, New York, United States Minor Outlying Island*<sup>3</sup> *IIED, Evaluation, London, United Kingdom*

This panel is designed to encourage participants to share and learn about the role that evaluation plays in relation to the 2030 Agenda for Sustainable development; also called the Sustainable Development Goals (SDGs). Almost all nations have signed the Agenda; evaluation is mentioned as an important part of follow-up and review but what part are evaluators playing? This session aims to build awareness of how evaluators can play an important role at any level in SDGs implementation and review and can contribute to the global transformation agenda. EVALSDGs – Evaluation – Value Added and Learning for the SDGs is a network of evaluation practitioners that advocate for evaluation practice and evaluative thinking to be an integral part of the SDGs implementation and review process. Since 2015 many actions have occurred and considerable experience has been achieved. However, time is moving fast and if evaluation is to play a critical role, evaluators worldwide need to become active in ensuring that evaluation is embedded in SDGs processes. This interactive session will initially present at least three case studies by EVALSDGs members of where evaluation has made a difference in the process of the SDGs implementation. Each of the three presenters will share their experiences in different countries in relation to the national processes for SDGs planning and implementation, and the role that evaluation is playing. The interactive session will encourage participants to share their own experience of engagement with the SDGs in their own country, whether positive, or frustratingly little. This will provide an opportunity for participants to share how they may be able to increase the action for evaluation in their own countries. The discussions will be documented and will be used to build knowledge on this important and increasingly critical area of engagement for the evaluation sector. A summary of the session will be generated and be used to further the work of advocacy for evaluation for the 2030 Agenda.

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**FBS 03 EVALSDGs - Evaluation Is an Essential Ingredient for the SDGs - Sharing Experiences**

## **FB 08 - Contribution of Dorothy Lucks**

*D. Lucks*<sup>1</sup>

<sup>1</sup> SDF Global, EVALSDGs, Perth, Australia

**FBS 03 EVALSDGs - Evaluation Is an Essential Ingredient for the SDGs - Sharing Experiences**

## **FB 09 - Contribution of Stefano D'Errico**

*S. D'Errico*<sup>3</sup>

<sup>1</sup> IIED, Evaluation, London, United Kingdom

**FBS 03 EVALSDGs - Evaluation Is an Essential Ingredient for the SDGs - Sharing Experiences**

## **FB 10 - Contribution of Kassem ElSaddik**

*K. El Saddik*<sup>1</sup>

<sup>1</sup> EVALSDGs, Lebanon

**FBS 03 EVALSDGs - Evaluation Is an Essential Ingredient for the SDGs - Sharing Experiences**

## **FB 11 - Contribution of Ada Ocampo**

*A. Ocampo*<sup>1</sup>

<sup>1</sup> UNICEF, Evaluation, New York, United States Minor Outlying Island

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

#### FBS 04 Evaluation Confidential - Adventures From The Evaluator's Kitchen

### Evaluation Confidential – Adventures from the Evaluator's Kitchen

*T. Kliest<sup>1</sup>*

<sup>1</sup> *Independent consultant, Houten, Netherlands*

How do we achieve evaluation quality within the parameters of time, scope and resources? An evaluation is an undertaking filled with expectations from all stakeholders involved, often competing expectations, and sometimes unrealistic in relation to the availability of information, stakeholders, primary and secondary data and allocated resources. How do we as evaluators manage these expectations while ensuring that the evaluation maximizes learning for the intended users and remains a positive experience for all involved?

Drawing from the experience of undertaking over 250 evaluations for a range of international development cooperation organisations in the past seven years, NIRAS' Evaluation Unit shares a behind-the-scenes look at the Evaluator's "kitchen" where things sometimes get hot while evaluations are planned and implemented. Both internal and external challenges and limitations will be reflected on to summarise the lessons we have gained from success and failure of evaluation processes, specifically:

- Implications of allocated evaluation resources – juggling between technical time versus management time and finding the right balance between time for familiarisation versus time for evaluation
- Approaches to realistic planning and expectation management – informing and involving stakeholders in evaluative activities while respecting schedules and expectations
- Logistics and practicalities of undertaking evaluations – organising realistic field work plans while accounting for uncertainty and preparing for the unexpected
- Mobilising evaluation teams and developing national evaluation capacities – engaging with team members to ensure a collaborative design and implementation of the evaluation and making the most of a mix of capabilities
- Data collection challenges – adapting data collection approaches to the conditions in which information can realistically be gathered while ensuring a respectful and fruitful engagement with respondents and stakeholders
- Stakeholder management – responding to stakeholders' requests and learning interests while managing expectations and ensuring the evaluator's independence

The objective of the session is to contribute to learning around the practicalities of designing and managing evaluations. If you are an evaluator you will recognise the challenges and we hope to encourage you to share your own experience, kitchen tips and recipes. If you're an evaluation commissioner or an evaluation stakeholder these insights may help you plan more realistic and useful evaluations.

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**FBS 04 Evaluation Confidential - Adventures From The Evaluator's Kitchen**

## **FB 12 - Introduction to NIRAS' Evaluation Approach, Process and Examples**

*C. Ljungman*<sup>1</sup>

<sup>1</sup> NIRAS Sweden AB, Evaluation & Results, Stockholm, Sweden

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

FBS 04 Evaluation Confidential - Adventures From The Evaluator's Kitchen

## FB 13 - Implications of Allocated Evaluation Resources – Juggling Between Technical Time Versus Management Time and Finding the Right Balance Between Time for Familiarisation Versus Time for Evaluation

*J. Rothman*<sup>1</sup>

<sup>1</sup> NIRAS Sweden AB, Evaluation & Results, Stockholm, Sweden

How do we achieve evaluation quality within the given parameters? Each evaluation is an individual “project” that must be undertaken within the triple constraints of time, resources and scope, often referred to as the project triangle. These three factors have to be well balanced to ensure high quality of the evaluation process and outputs, and stakeholders must share a common understanding of the specific constraints in order to reach agreed expectations.

Drawing from the experience of undertaking over 250 evaluations for a range of international development cooperation organisations in the past seven years, this presentation specifically discusses NIRAS' experience of planning and managing evaluations, specifically balancing the scope with the available resources in order to create realistic expectations and high quality outcomes. We present our lessons learned in relation to:

**Resources:** the allocated evaluation resources (budget) must match the scope of the evaluation, or trade-offs have to be made. Travel costs related to data collection must be sufficiently covered and allocated time (working days) must be realistic in relation to the work that is needed to be carried out to achieve the expected scope. One of the most common limitations highlighted in evaluation reports relate to inadequate resources for primary data collection in the field and data analysis.

**Scope:** the scope must be clear, realistic and jointly understood by all the stakeholders. Agreeing on the scope includes defining what will be evaluated, which evaluation criteria and key evaluation questions to include. Availability of secondary data, access to data sources, data reliability and stakeholder engagement are important factor to consider when defining the scope in relation to the allocated time and resources.

**Time:** The availability of time is often one of the most critical factors for ensuring evaluation quality. The time constraint can be broken up into three aspects:

- a) *Lead time* – the time between the tendering, contracting and start-up of the evaluation needs to be realistic in order to ensure availability of all stakeholders;
- b) *Calendar time* – the timeframe for the evaluation must allow sufficient room for stakeholders and informants to engage, and the evaluators must have time to undertake the required work. Many evaluations are squeezed into the minimum duration possible and inception phases are often inadequate which has a direct result on evaluation quality;
- c) *Black-out periods:* An understanding of the periods that are not conducive to undertaking evaluations must be taken into account which include seasons, national events such as elections, or other internal processes which are not conducive for engaging stakeholders in evaluations. The availability of all stakeholders must be confirmed during the planning stage, so that roles and responsibilities can be defined and all can be held accountable for contributing to a constructive process.

We will present examples drawing from the 250 evaluations undertaken by NIRAS' Evaluation Unit to illustrate the main points above. This session is expected to contribute to a better understanding of these three factors (time, resources and scope) to help commissioners create the pre-conditions for high quality evaluation processes.

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

FBS 04 Evaluation Confidential - Adventures From The Evaluator's Kitchen

## FB 14 - Stakeholders Management – Responding to Stakeholders' Requests and Learning Interests while Managing Expectations and Ensuring the Evaluator's Independence

*R. Bisiaux*<sup>1</sup>

<sup>1</sup> NIRAS Sweden AB, Evaluation & Results, Stockholm, Sweden

An evaluation is an undertaking filled with expectations from all stakeholders involved, often competing expectations, and sometimes unrealistic in relation to the availability of information, stakeholders, primary and secondary data and allocated resources. Drawing from the experience of undertaking over 250 evaluations for a range of international development cooperation organisations in the past seven years, this presentation specifically discusses NIRAS' experience of stakeholders management and the importance of responding to learning needs while ensuring the independence of the evaluator. We present our lessons learned in relation to:

- **Understanding the politics of the project or intervention being evaluated:** Relevant stakeholders should be involved early in defining the rationale and purpose of the evaluation. To ensure that key groups and their respective perspectives are incorporated at appropriate points throughout the evaluation, it is important for the evaluator to plan ahead which stakeholders should be engaged and what their role in the evaluation should be. Who are the relevant stakeholders? What are their stakes in the evaluation? Should agreement or resistance be expected from them? And between the evaluator and the stakeholders?
- **Clarifying expectations:** To resolve any disagreements or misunderstandings about the overall purpose or specific evaluation questions among stakeholders, the evaluator needs to consider how best to communicate but also engage in a dialogue with the different stakeholders. A consultative process will help build shared ownership around the evaluation, a clear understanding of what is and what isn't being evaluated, and this will eventually facilitate the reception of findings in later stages.
- **Managing diverging learning interests and requests:** Stakeholders often have conflicting requests reflecting their various interests and stakes in an evaluation. Unrealistic expectations, out-of-scope requests and strong steers might derail the evaluation from its path. However, the needs and constraints of stakeholders should be accounted for with a view to increase the utility of the evaluation to its users. It is also possible that the specific learning needs in relation to the project being evaluated only emerge after the evaluation has been commissioned, which suggests that in some cases the scope and focus of the evaluation might change to reflect this and increase the utility of the evaluation.
- **Ensuring independence:** Maintaining independence as evaluators as we get caught up in the middle of conflicting requests and stakeholders can be a complex endeavour. Flexibility, anticipation and clarity over the role of the evaluator are key, to contribute to the credibility of the evaluation.

We will present examples drawing from the 250 evaluations undertaken by NIRAS' Evaluation Unit to illustrate the main points above.

This session is expected to contribute to a better understanding of the benefits and challenges of involving stakeholders in the evaluation process. It should influence evaluators to better plan their strategy for engaging stakeholders, managing their expectations and responding to their learning interests throughout the evaluation cycle. Other evaluation stakeholders may also get clarity over their roles in an evaluation.

Friday, 5 October 2018  
13:45–15:15

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**FBS 05 The Work of the European Commission Regulatory Scrutiny Board**

## **The Work of the European Commission Regulatory Scrutiny Board**

*B. Naudts*<sup>1</sup>

<sup>1</sup> European Commission, Regulatory Scrutiny Board, Brussels, Belgium

This Fishbowl session will build on the discussion in Flagship symposium 4. Its main purpose is to discuss further those issues that could not be delved into deeply enough in the Flagship session.

**FBS 05 The Work of the European Commission Regulatory Scrutiny Board**

## **FB 15 - Contribution of Vesna Valant**

*V. Valant*<sup>1</sup>

<sup>1</sup> Slovenia

**FBS 05 The Work of the European Commission Regulatory Scrutiny Board**

## **FB 16 - Contribution of Joakim Nilsson**

*J. Nilsson*<sup>1</sup>

<sup>1</sup> Sweden

**FBS 05 The Work of the European Commission Regulatory Scrutiny Board**

## **FB 17 - Contribution of Michael White**

*M. White*

**FBS 05 The Work of the European Commission Regulatory Scrutiny Board**

## **FB 18 - Contribution of Guenther Ebling**

*G. Ebling*<sup>1</sup>

<sup>1</sup> European Commission, Brussels, Belgium

# Special Session

Thursday, 4 October 2018  
11:30–13:00

## SS 04 Panel on Regional VOPEs

Z. Mousa<sup>1</sup>

<sup>1</sup> Lebanon

The conference takes the opportunity of the presence at EES2018 of several regional evaluation associations' presidents to debate issues of common interest. Alongside national VOPEs – of which there are many, e.g. in Europe nearly all countries have one – it is important that regional evaluation associations add specific value to the evaluation eco-system. So, what is their specific role? What can they offer that national and other level societies are less well placed for? What is their additionality? How can regional and national VOPEs better benefit from each other? Can good practice on how to best work together at different levels be identified? Should regional VOPEs cooperate and exchange more and better between themselves to reach their goals? And how?

These issues will be discussed by a Panel composed of Adeline Sibanda (President, AfrEA), Leslie Goodyear (President, AEA), Tessie Catsambas (President Elect, AEA), Silvia Salinas (Coordinator, RELAC) and Bastiaan de Laat (President, EES). The Panel will be moderated by Ziad Mousa (IOCE, EvalPartners) and the current NESE Chairman, IOCE member and EES Board member Weronika Felcis will act as discussant. After brief introductory statements of each of the panel members and the discussant, the discussion within the Panel, and in particular with the audience, will be open.

## SC 03 – Panel Contribution

A. Sibanda<sup>1,2</sup>

<sup>1</sup> AFREA, IOCE, President, Johannesburg, Zimbabwe

<sup>2</sup> S2SE, chair

## SC 04 – Panel Contribution

A. Tzavaras Catsambas<sup>1</sup>

<sup>1</sup> EnCompass, Evaluation, Rockville, USA

## SC 05 - Panel Contribution

B. de Laat<sup>1</sup>

<sup>1</sup> European Evaluation Society, Board, Paris, France

## SC 06 - Panel Contribution

L. Goodyear

## SC 07 - Panel Contribution

S. Salinas<sup>1</sup>

<sup>1</sup> Independent Evaluator, Coordinator RELAC, La Paz, Bolivia

## SC 08 - Panel Contribution

W. Felcis<sup>1</sup>

<sup>1</sup> Freelance evaluator, Riga, Latvia

# Paper Sessions

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 001 Approaches for and Consequences of Big Data****O 001 - Approaches for Making Use of Unstructured Big Data in Evaluation**Y. Ofek<sup>1</sup><sup>1</sup> *University of Haifa, Public Management and Policy, Haifa, Israel*

The amount of electronic data generated around the world is constantly growing, and there is no doubt that we will have to deal with more and more of it (Mayer-Schönberger & Cukier, 2013). Electronic devices, computer systems, websites, and information systems produce vast amounts of both structured and unstructured data.[1] While the literature is replete with examples of using structured Big Data for evaluation (and subsequently decision making), systems that can analyze unstructured Big Data for evaluations, including texts, are rarer.

Currently, evaluators approach textual Big Data through traditional content analysis methods based on keyword search (Marz & Warren, 2015; Törnberg, 2016; Wiedemann, 2013). This is a valid and important approach, which is likely to be used also in the future. However, this practice is also limited to iterating over predefined concepts.

In Big Data, where petabytes of text are intertwined, not only that the number of keyword combinations is enormous, but sometimes our goal as evaluators cannot be defined using keywords. To overcome this problem, some researches suggest automated methods to analyze big data. In a recent study entitled "Automatic Content Analysis Methods for Political Texts," Grimmer and Stewart (2013) suggest overcoming these limitations by providing detailed guidance for using automated methods as a massive cost reduction approach to analyzing even moderately sized collections of texts. The most innovative systems, combine machine learning together with visualization techniques, providing evaluators with the ability to reveal trends in big data – a capacity they did not have before (Bekkerman & Raveh-Robicek, 2017).

This paper and presentation will review the various main approaches of using big data for evaluations and by evaluators, who should not have prior technical knowledge in computer science or big data. The disadvantages and advantages of the approaches will be reviewed, focusing on when and how evaluators can use them, and for which type of evaluation. This review will demonstrate the use of these methods by focusing on different case studies.

[Foot Note 1] Structured data are generated manually or automatically by following a specific construction pattern called schema. Any dataset that can be stored as a table with rows corresponding to data instances and columns corresponding to their attributes is considered structured. In contrast, unstructured data is free of schema. Documents, photographs, and videos are all forms of unstructured data. An example of structured data would be a list of high school students with their names, addresses, and grades, whereas their essays would be unstructured data.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 001 Approaches for and Consequences of Big Data**

**O 002 - Can the Use of Big Data Overcome the Limits of Large Scale Surveys and Other Types of Stakeholder Consultations?**

*C. Moeller*<sup>1</sup>

<sup>1</sup> *Optimity Advisors, Public Policy Practice, Berlin, Germany*

Collecting relevant and accurate data is at the heart of every evaluation and crucial in reaching robust conclusions. The most common data collection tools involving stakeholders are surveys, interview programmes, and focus groups/workshops. Usually subject matter experts benefit from workshops/focus groups due to their interest in evaluation outputs and due to visibility/networking effects. It can however be more difficult to gain relevant information through surveys and interviews due to fatigue and time limitations of stakeholders. There is thus a need to rethink ways to gather large sets of data particularly by minding opportunities posed by ICT tools. Respectively, the paper will discuss the use of social media data and big data as alternative (or additional) elements to involving large numbers of stakeholders to collect data. By elaborating on two evaluations as case studies, the paper examines on the advantages and challenges of this data collection method. The first case study is an evaluation of an EU agency where large web datasets have been analysed to gather views on the media's sentiments towards FRA. The second case study is an evaluation of an UN awareness raising campaign where social media data was used to analyse the usage of social media of the target group. By leveraging on experiences from both case studies, several advantages of using big data will be discussed such as: (i) the accurate and holistic nature of the data, (ii) the time and cost-efficient access to relevant and up-to-date data, and (iii) overcoming survey fatigue and/or lack of participation due to confidentiality concerns. In addition, key challenge will be analysed such as data protection concerns in case that data is not anonymised and lack of contextual information and thus concerns in relation to wrong assumptions. Respectively, the paper concludes that Information Communication Tools have created new data sources such as big data which could overcome challenges posed by traditional data collection tools such as lack of stakeholder participation, incorrect or outdated data and rigidity by allowing for agility and the access to large data sets which provide a precise overview of locations, sentiments and usage. At the same time, it will be important to understand how synergies between new data sources such as big data and traditional data sources can be generated to guarantee best results.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 001 Approaches for and Consequences of Big Data**

**O 003 - Collateral Damages by Shotgun Shooting: Why Big Data May Lead to Big Trouble**

*W. Meyer<sup>1</sup>*

<sup>1</sup> Saarland University, CEval, Saarbrücken, Germany

Progress in IT and internet communication offers a lot of new opportunities for evaluations. The pure size of information attracts many people, intensified by the simpleness of its use. The use of some applications increased enormously, like for instance search machines for hotels or restaurants including evaluation tools. One may even think about replacing evaluations by fully automated big data applications. The fascination of modern technologies may obstruct someones view on the shortcomings of big data. This paper discusses the implications of big data and the pitfalls of its use in evaluations. It is possible to distinguish three different processes with specific effects and risks. First, one has to discuss the acceptability of data collection through data or text mining software. Key areas of discussion are the identification of basic population, the adequacy of selection procedures and the way of clustering results. Second, big data is about new analytical tools and algorithms. There are different advanced and very complex concepts and solutions, so it is not easy to discuss the pros and cons of these approaches. Finally, big data solutions deliver recommendations and decisions. The quality of these results and the evaluation criteria used are another important area for investigations. The paper will roughly go through these aspects and discuss how they can be controlled and used for evaluations.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 001 Approaches for and Consequences of Big Data

## O 004 - Is it Possible to use Big Data for UN Women Evaluations?

S. Reddy<sup>1</sup>, C. Lopes<sup>2</sup>

<sup>1</sup> UN Women, Evaluation, New York, USA

<sup>2</sup> Cambridge University, Social Science, Cambridge, United Kingdom

There is a question mark around the feasibility of leveraging big data and data analytical techniques in relevant, ethical and efficient ways to improve evaluation.

This presentation will focus on a UN Women commissioned study in 2017 to determine to what extent it might be able to strengthen traditional evaluations using additional evidence streams from big data, with a focus on piloting social networking, media and news platforms and radio data.

Taking Mexico and Pakistan as two case studies that present different challenges to access of big data sources, and distinct barriers to women's leadership and political participation, we will discuss the process of accessing, analysing and triangulating big data with other data sources. The findings highlight that social media data should be seen as evaluative, but due to its bias, omissions and observational nature, it should not be analysed in isolation. Big data should be considered as another source of rich data, which is more connected to individuals and therefore generalizable across life contexts, that complements the findings of evaluations based on traditional data.

The results support a better understanding of how UN Women and its partners might effectively use big data for evaluation efforts and raise additional questions about its use that requires further exploration. This includes supporting efforts to "Leave No One Behind" in evaluation, with new possibilities for enhancing the voice and power of marginalized or hard-to-reach groups within evaluations. The commissioners, authors and members of the Advisory Group will present the study and discuss the implications of its recommendations for future work on big data and evaluation.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 002 Evaluating the Sustainable Development Goals

## O 005 - How Firms Evaluate Their Contribution to the Sustainable Development Goals: Tools, Gaps and Future Challenges

N. Schönherr<sup>1</sup>, A. Martinuzzi<sup>1</sup>

<sup>1</sup> Vienna University for Economics and Business, Institute for Managing Sustainability, Vienna, Austria

Upon the endorsement of the Sustainable Development Goals (SDGs) in 2015, Ban Ki-moon, former Secretary-General of the United Nations, called them a 'paradigm shift for people and the planet'. Part of this paradigm shift is the assertion that businesses, governments and civil society actors are equally called upon to pursue this agenda. Evaluating the contribution of business to the SDGs poses particular challenges. For firms, it requires a well-founded understanding of the wider impacts they cause through their core business, community and philanthropic engagement, as well as the materiality of these impacts in a sustainable development context. For the wider evaluation community, it implies a need to leverage the tools and practices of firms in measuring impacts on sustainable development for establishing the value, worth and merit of the business contribution to the SDGs.

This presentation critically analyzes the respective qualities of selected tools for measuring sustainable development impacts of firms. In doing so, it elicits gaps and future challenges for high-quality impact measurement and management in a corporate context and draws conclusions for the evaluation of the business contribution to the SDGs.

Based on an analysis of 200+ tools, we identify four significant gaps: First, extant tools and corporate practice frequently focus on internal processes and activities, at the expense of indirect and collective impacts accruing from them. Second, they are prone to selective boundary setting and scoping with a strong focus on physical transactions and lower priority accorded to relational aspects (e.g. impacts on stakeholders, communities and wider society). Third, they tend to disregard the growing body of data on changes in the socio-environmental systems in which they are embedded (e.g. data on sustainability trends, risks and opportunities). Finally, firms do not always meet basic standards of information reliability and completeness (e.g. independent verification or reliable indicators).

These gaps lead us to formulate a number of important challenges and discuss first steps towards addressing them for improving our understanding of the business contribution to the SDGs. Specifically, we highlight the need to improve extant corporate tools and practice to enhance their capacity to support systematic mapping and measuring of impacts in the context of the SDGs. We argue that corporate measurement and evaluation systems will need to (a) carefully delimit the purpose and ambition of measurement; (b) take explicit account of environmental and social system boundaries; (c) align corporate Key Performance Indicators (KPIs) with societal indicators apt to capture multidimensional and systemic effects of corporate activities on society and fill corresponding data gaps, and (d) develop and openly discuss with stakeholders the (e) valuation factors that determine how results are valued, prioritized and translated into action.

Improved impact measurement and evaluation provides a real opportunity to glean more substantial information on the business contribution to the SDGs. At the same time, a deeper understanding, better data and a more transparent approach to evaluation can help firms regain public trust and fulfil their role as an important partner in the achievement of the SDGs.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 002 Evaluating the Sustainable Development Goals

## O 006 - Global Evaluation Agenda 2016–2020, Bite of EvalApple?

*A. Kalugampitiya*<sup>1</sup>

<sup>1</sup> EvalPartners, EvalPartners, Colombo 5, Sri Lanka

In 2014, EvalPartners started a global, multi-stakeholder consultative process to brainstorm about the priorities and key areas of a Global Evaluation Agenda for 2016–2020, which is called “EvalVision2020.” In 2015, this agenda-setting consultation continued face-to-face in over 92 global, regional and national EvalYear events, with each event invited to contribute additional ideas to EvalAgenda2020. The Global Evaluation Agenda was launched at the penultimate EvalYear Global Event held in Kathmandu, Nepal at the parliament of Nepal.

The Global Evaluation Agenda is important in many ways: *its participatory nature* as the Global Evaluation Agenda was developed in consultation with the global evaluation community. A months of online consultation on top of face to face dialogues around the world ensured valuable inputs and participation of professionals all corners in the globe: *ownership of the agenda* as it's owned by everyone contributed to its development and anyone interested to take part in implementation. *It's nature for anyone to contribute* as any individual, any organization, any government or interested party can contribute to any parts of the Global Evaluation Agenda by implementing relevant initiatives in respective geographies.

The consultation for EvalAgenda2020 has shown that evaluation, in order to reach its fullest potential, must combine effective methods and techniques and the values that drive policies geared to the public interest. That is, we collectively support evaluation as a value-driven tool for improved policy-making, governance, program design, program implementation and ultimately, to achieve outcomes that are more equitable, inclusive and sustainable for all people. And we are aware that in order to achieve such expectations we need to focus on both the demand and supply dimensions of the evaluation process.

How can evaluation help to achieve this dream? Our vision for 2020 is that evaluation is an integral part of all efforts by governments, civil society, and the private sector to improve the lives and conditions of all citizens. Our vision is that high-quality and value-driven evaluation can improve the design and implementation of these efforts, track their progress, make mid-course corrections and assess final outcomes and impacts with a view to social learning across policies, programs and initiatives. This paper is to highlight key components and key achievements of the Agenda2020 (<https://evalpartners.org/global-evaluation-agenda>)

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 002 Evaluating the Sustainable Development Goals**

**O 007 - Strengthening Institutions and National Evaluation Capacity for Achieving the SDGs: Increased Accountability, Transparency, and Public Policy Evaluation in Morocco**

C.A. Asenjo Ruiz<sup>1</sup>, K. Kayser<sup>1</sup>

<sup>1</sup> *United Nations Office on Drugs and Crime, Independent Evaluation Unit, Vienna, Austria*

The session aims at discussing the important role of accountability, transparency and evaluation of public policy for achieving the Sustainable Development Goals (SDGs) and, in particular, Goal 16 which aims at “promoting peaceful and inclusive societies for sustainable development, providing access to justice for all and building effective, accountable and inclusive institutions at all levels”. The presentation will focus on the paramount role of public policy evaluation for enabling evidence-based decision making in the public sector. International efforts towards more accountable and transparent institutions will be showcased by sharing the experience of the Kingdom of Morocco since the Constitutional reform of 2011, which established the mandate to evaluate all public policies as part of the oversight role of the Parliament. In the session, it will also be presented the new Master’s programme for Public Policy Evaluation of the University Mulay Ismail of Meknes, the first of its kind in Morocco, which was launched in October 2017. The Master’s programme has been developed by the University Mulay Ismail in close collaboration with Morocco’s National Observatory for Human Development and support of the UNODC’s Independent Evaluation Unit.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 003 Better Regulation Of EU Policies And Programmes

## O 008 - Is Better Regulation any Good for European Citizens?

*L. Luca*<sup>1</sup>

<sup>1</sup> *European Commission, DG TAXUD Taxation and Customs Union, Brussels, Belgium*

The European Union and its institutions have been going through turbulent times and have been under increased scrutiny in the last years over their policies and even their rationale for existing. Citizens find it difficult to understand the complex institutional structure and the ways in which EU rules are created and then implemented.

In order to address some of the criticism, the European Commission has put forward in 2015 the Better Regulation Guidelines. These guidelines represent a set of rules and best practices that define how retrospective evaluations and impact evaluations are to be carried.

The objectives were to deliver high-quality Union legislation which is efficient, effective, simple, clear and which avoids over-regulation and administrative burdens for citizens, public authorities and businesses. More specifically, the aim of the guidelines was to open up the policy making process to regular citizens, involve them to a higher degree from the start of an initiative, increase the transparency of the overall process and allow for a final check before the EU rules become law. To give more access to citizens which have so far not been able to participate in the way the decision have been taken, going beyond the regular actors which have so far shaped the debate, especially the lobbying groups.

The purpose of my presentation will be to check whether these objectives have been met in practice. I plan to use real life examples from the different EU initiatives I have worked on (for example the Tobacco Directive, the Cash Controls Directive or the Digital taxation proposals) and see to what extent policy making has taken into account the opinion of stakeholders and their influence on it. I will also look at the different public consultations that have accompanied each of these initiatives and test whether they have successfully managed to reach out to new stakeholders or have only been limited to engaging with the regular actors.

I will conclude with the lessons learned so far from the guidelines over the 3 years of implementation and possible ways to improve the current rules in order to better reach their aims.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 003 Better Regulation Of EU Policies And Programmes

## O 009 - Evaluation of an EU Funding Programme Which, Due Turbulent Times, Took a Very Different Form

*V. Ludden*<sup>1</sup>

<sup>1</sup> *Ramboll Management Consulting, European Policy, Brussels, Belgium*

When the number of migrants and asylum seekers rapidly increased through 2015 and 2016, it put a severe strain on the asylum and reception systems of several Member States of the European Union. Not least the Asylum, Migration and Integration Fund (AMIF) was tested, dedicated to promoting the efficient management of migration flows across Europe and to the implementation, strengthening and development of a common Union approach to asylum and immigration. As the Fund was set up during the pre-migration crisis period, i.e. within a very different context to that which it ultimately had to operate in, the interim evaluation that Ramboll Management Consulting conducted for the European Commission in 2017 and 2018 proved important not only in terms of assessing the performance of the Fund at its half-way mark, but also with regard to planning ahead, post 2020, and assessing future options to enable the European Commission to support Member States should a crisis of this magnitude re-occur.

The evaluation applied a theory-based approach and a mix of methods, ranging from the meta-evaluation of Member States' interim evaluation reports, to desk research, to in-depth interviews with key stakeholders, to case studies in a selection of seven countries. The evaluation team faced a number of design and management challenges when carrying out the study, including:

- Applying the European Commission's Better Regulation Guidelines and pre-defined evaluation questions as per AMIF's legal acts in a dramatically different context, highlighting the at times prescriptive nature of such evaluations, which can be less adaptable to changing contexts and open to new questions.
- Carrying out an interim evaluation of a programme which was strongly influenced by external factors and consequently took a form that was very different to what was originally foreseen, while still having to report in relation to set targets and baselines. Within this framework it proved important to balance the requirement to report on pre-defined quantitative indicators with the need to properly contextualise them to come up with valid and fair conclusions.
- Implementing the evaluation of a wide-ranging, multi-faceted programme with the help of a variety of tools within a very short timeframe due to the European Commission's need for results to inform its plans for the next programming period. The meta-evaluation of the national interim evaluation reports was carried out within the space of one month which brought with it a number of management challenges.

The evaluation therefore helps shed light on the challenge of evaluating, within the constraints of a pre-defined framework, the performance of an EU-level programme which fell prey to external, turbulent factors. Ultimately the evaluation created an interesting overview of interventions across borders and represents a potential knowledge source for national development. This overview gave rise to a set of recommendations for future actions regarding the design, implementation, management and monitoring and evaluation of AMIF and future Funds.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 003 Better Regulation Of EU Policies And Programmes

## O 010 - Evaluating Gender Mainstreaming in European Research & Innovation Policy: Interim Evaluation of Gender as a Cross-Cutting Issue in Horizon 2020

*M. Bustelo<sup>1</sup>, B. Knoll<sup>2</sup>, S. de Cheveigné<sup>3</sup>, E. Engebretsen<sup>4</sup>, U. Sandström<sup>5</sup>*

<sup>1</sup> Complutense University of Madrid, Dpto. C. Política y de la Admón. II. Facultad de CC.Políticas y Sociología, Madrid, Spain

<sup>2</sup> B-NK GmbH, Consultancy for Sustainable Competence, Vienna, Austria

<sup>3</sup> French National Centre for Scientific Research CNRS, Centre Norbert Elias, Marseille, France

<sup>4</sup> University of Oslo, School of Medicine, Oslo, Norway

<sup>5</sup> School of Industrial Engineering and Management at KTH, Department of Industrial Engineering and Management, Stockholm, Sweden

The evaluation of gender mainstreaming strategy and policies has been recognized to be a challenging and complex endeavor. Some methodological problems associated with treating gender mainstreaming as a goal rather than as a means to the long-term objective of gender equality have been recognized (Hunt & Brouwers, 2003). The lack of identifying intermediate results for that long-term objective has been also pointed out (Bustelo, 2003).

Although gender mainstreaming implementation has not been a success across all areas, EU research policy provides a more positive example of gender mainstreaming (Hafner-Burton & Pollack, 2009). It was on the 5<sup>th</sup> Framework Programme (FP5, 1998–2002) that gender was included through gender equality objectives within the legal basis underpinning EU research policy. The EU Research & Innovation Policy has had an important evaluation activity around their Framework Programmes, being specifically important the Interim and Ex-post evaluations, at least since FP5. Until FP7, the intermediate and ex-post evaluations were conducted by a "High Level Expert Panels" which include in their evaluations all Programmes aspects, including cross-cutting issues. However, as part of a new evaluation policy implemented, for the intermediate evaluation of Horizon 2020 (8FP), several independent evaluations of cross-cutting issues were commissioned by D.G Research for feeding the general intermediate evaluation. Among them, the one on gender equality, which was conducted between November 2016 until April 2017.

This paper aims at presenting the evaluation done following this new evaluation approach, as well as the methodological challenges to evaluate the three cross-cutting gender equality objectives in Horizon 2020: gender balance in research teams at all levels, gender balance in decision making, and integrating the gender dimension in the content of R&I. The paper presents the methodological approach followed by the evaluation study, and the main limitations found due mainly to lack of time and reliable data, as well as some important results, which might be especially significant for a successful gender mainstreaming strategy, such as the importance of gender expertise and gender training to key stakeholders. Also, the paper concludes with some lessons learned and some recommendations for the evaluation of crosscutting issues in EU policies.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

#### S 004 Resilient Climate Change

### O 011 - Using Evaluation to Understand the Potential to Scale Up Resilience Innovations

*C. Stott*<sup>1</sup>

<sup>1</sup> *Itad, Climate Change theme, Hove, United Kingdom*

The Global Resilience Partnership (GRP) is an independent partnership funded and engaged in by public and private organisations working towards a shared vision of resilience for vulnerable people and places. In its first phase, from 2015–2017, GRP supported the implementation of two challenge rounds. Grants of up to USD 1 million, provided by USAID and the Z Zurich Foundation, were delivered to 23 selected grantees throughout Africa and Asia to implement innovative solutions for resilience building. With initial implementation now tested, evaluation can explore which of these innovations have potential to be scaled up and the potential pathways through which this can be done.

Resilience building in the context of GRP brings together multi-stakeholder partnerships and combined solutions, to enhance resilience of the most poor and vulnerable to multi-faceted risks, such as climate change, ecosystem fragility, disease outbreaks and geopolitical instability. Scaling of these models is therefore inherently complex and presents the unpredictability of introducing innovations to new and changing contexts.

Itad are conducting a formative evaluation of the GRP to understand how GRP's approach to managing these challenges, providing financial and technical support, and encouraging knowledge sharing among grantees, can support effective implementation; identify successes and challenges in implementing resilience innovations; and, consider how these efforts may be scaled up.

**We will share findings of how GRP's challenge innovations have been successfully introduced among vulnerable communities, the successes and challenges that grantees have faced in implementing them, and what this means for resilience building. We will discuss the ways in which these resilience innovations are showing potential to be scaled up to broader levels of implementation; or out to new social, political, geographical and ecological contexts.**

Findings of this evaluation have wide-reaching implications for how local action towards resilience can contribute to wider scale impact pathways for global resilience building. By setting their progress within the context of GRP's impact pathways, we will infer what this means for resilience building more broadly, to the benefit of those working to support vulnerable communities to adapt better to sudden shocks and chronic stresses. Finally, we will reflect on the role and contribution of the formative evaluation for resilience programmes operating in complex and changing contexts.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

#### S 004 Resilient Climate Change

### O 012 - The Role of Monitoring and Evaluation for Global Climate Resilience Programmes: Challenges, Lessons and Ways Forward

*P. Silva Villanueva*<sup>1</sup>, *V. Sword-Daniels*<sup>2</sup>

<sup>1</sup> *ResilienceMonitor, Director ResilienceMonitor, Madrid, Spain*

<sup>2</sup> *Itad, Climate Change, Brighton, United Kingdom*

**Rationale:** As development initiatives seek to address the realities of changing climates there is a pressing need to track progress and understand what works in building climate resilience, what does not, and why. The context-specific nature of climate resilience requires that monitoring frameworks are flexible enough to accommodate diverse geographies, contexts and scales, particularly when large programmes, funds and agencies work across these. In rapidly changing contexts there is an increased need to understand how resilience is being built during implementation, to support learning within the lifetime of a programme rather than after the fact. This understanding places renewed emphasis on the role and value of monitoring and evaluation systems to support learning within resilience building programmes.

**Objectives:** This paper presents the practical experiences of designing and implementing the monitoring and evaluation system for DFID's £140 million Building Resilience and Adaptation to Climate Extremes and Disasters (BRACED) programme to meet the dual objectives of accountability and learning. The BRACED programme, which operates across 13 countries in East Africa, the Sahel and Asia, is the biggest global effort to build resilience at a local level, in highly vulnerable places, yet at scale. The diversity of the portfolio poses challenges for the design of monitoring and evaluation systems to ensure quality and utility of evidence.

**Findings:** This paper reflects on the past three years of implementation, regarding the challenges and lessons learnt from developing, testing as well as refining a variety of resilience measurement and monitoring approaches and frameworks. These insights from practice can inform the development of global climate resilience programmes. The paper focuses on three methodological challenges related to monitoring and evaluation systems for programmes operating at scale through investments in large portfolios: i) dealing with the trade-offs of programme level monitoring efforts (accountability versus learning, quantitative versus qualitative information, programme versus project level learning needs, internal versus external audiences and expectations); ii) adapting programme M&E frameworks and monitoring systems as projects evolve and mature; and iii) measuring and reporting resilience outcomes. This paper pays particular attention to the on-going project-to-programme real time learning and implications for adaptive programming and management.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

#### S 004 Resilient Climate Change

### O 013 - Coping with The Tragedy of the All Too Common: Common Themes in Evaluating Climate Change Mitigation

*R. Warne*<sup>1</sup>, *S. Chaplowe*<sup>2</sup>, *A. Cleaveland*<sup>3</sup>

<sup>1</sup> *ClimateWorks Foundation, Evaluation and Learning, San Francisco, USA*

<sup>2</sup> *Children's Investment Fund Foundation, Evidence- Monitoring and Evaluation, London, United Kingdom*

<sup>3</sup> *ClimateWorks Foundation, Strategy and Planning, San Francisco, USA*

**Rationale:** Our societies' ability to mitigate and be resilient to climate change is a crucial challenge. Climate change mitigation is a relatively new and growing field of work. It is crucial that we learn quickly what does and does not work and in what circumstances – both because of the urgency of shifting the trajectory of climate change and because of the need to be effective within and across complex, sometimes unpredictable political climates. As yet, there are few tried and true best practices for climate change work – so climate change projects are often necessarily breaking new(ish) ground with theories of change, implementation strategies and evaluation designs. Evaluations must successfully navigate the critical need for learning and feeding frank discussions about what is and is not working, and why, with the vital importance of doing no harm to efforts to reduce greenhouse gas emissions. Quickly climbing the steep learning curve is imperative.

**Objectives:** Organizations that do climate work are definitionally working in turbulent times, and our evaluations must take this into account by providing clear, real-time learning that enables current and future work to be effectively designed and implemented. To enhance learning among program managers and grantees – and to improve our own evaluation systems and approaches – evaluation staff from ClimateWorks Foundation, Children's Investment Fund Foundation, Hewlett Foundation, and MacArthur Foundation examined nine recent climate change mitigation evaluations that we and partners commissioned in the US and internationally (including China, Brazil, and India). We hope that what we learned can generate discussion around whether the themes we identified resonate in others' work and, if so, how to take these themes into account in evaluation.

**Content:** This paper lays out 1) the common themes that emerged across the evaluations and why we think they are so common; 2) methods we used to support foundation staff to explore applying lessons from these themes; and 3) how attention to common themes can help guide program design and evaluation work. Across the nine evaluations, the most common themes we found evaluators raising included

1. Importance of well-developed theories of change
2. Balance between feasibility and ambition of strategies and grants
3. Focus on mitigation vs. "mitigation plus" co-benefits
4. Funder coordination
5. Importance of individual leadership capacities and networks

While these themes were generated from climate change evaluations, most were familiar from evaluations we have done in other sectors. This raised the question – why are we seeing these themes come up so frequently? As part of addressing this – and to facilitate evaluation utilization – the authors led discussions with program staff from a variety of funders to explore why these themes were common, and we will present a summary of what they identified as means to build on these evaluation findings to enhance current and future programming. We will summarize ideas from the literature that help explain why these themes are so pervasive. We will also present what we have learned about building attention to these themes into other evaluative work.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 005 Institutionalising Evaluation

## O 015 - Challenges in Institutional Set-Up After Devolution from Federal to Provincial Government – A Unique Case of PWD

*S.A. Ali*<sup>1</sup>

<sup>1</sup> Pakistan Evaluation Association PEA, Board Member, Wah cantt, Pakistan

This evaluative study is a classic example of institutionalization that emerged in wake of the 18<sup>th</sup> amendment in the Constitution of Pakistan. Population Welfare Training Institute (PWTI) was transferred to the province from federal government after devolution. PWTI was established as an epic institution with the aim to develop capacity of Population Welfare Department (PWD) staff and officials, primarily through training programs and establishing a state-of-the-art research center, and repository. An evaluation of PWTI was conducted to assess what institution has achieved, and how it can be integrated in the newly formed provincial setup.

The objective of evaluation was “to determine the institution’s contribution towards effective capacity building initiatives since devolution”. A mixed method methodology was employed, initially a comprehensive literature review of PWTI, PWD, and documents related to 18<sup>th</sup> constitutional amendment was carried out, and a semi-structured “interview guide” was developed and administered to collect and analyse qualitative data from the staff and senior administration. Sample size was determined once the saturation arrived at after twelve in person interviews. A separate questionnaire was developed to validate beneficiaries – participants trained through capacity building initiatives and training programs. This questionnaire was administered through telephonic calls on 5% beneficiaries.

Analysis of quantitative and qualitative data highlighted certain gaps, which are; objectives of PWTI and PWD were not strategically linked, there was no proper institutional administrative structure, no defined hierarchy, and inadequate means for intellectual growth of staff through research. Subsequently, no mechanism devised for Monitoring & Evaluation (M&E) resulted in non-availability of data on capacity building initiatives implemented by PWTI. No pre/post training assessment, and feedback mechanism for beneficiaries were the other grey areas identified after data analysis.

As a result of evaluation, a precedent was established i.e., erection of a new institution legally and constitutionally to be followed by other devolved institutions. Also, the recommendations framed are; there should be a strategic fit among PWTI and PWD objectives, a robust results-based M&E mechanism along with a comprehensive Management Information System (MIS), and an active feedback mechanism should be developed. Indigenous innovative solutions proposed in evaluation may be adopted for smooth functioning of the new setup in the longer run ensuring sustainability and effective service delivery at the sub national level.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 006 Mixing Methods in Systematic Reviews**

**O 016 - Systematic Integration of Meta-Evaluation and Evaluation Synthesis: Aggregating Evidence Through a Quantitative Evaluation Synthesis Design**

*M. Noltze<sup>1</sup>, I. Verspohl<sup>1</sup>, S. Harten<sup>1</sup>*

<sup>1</sup> *German Institute for Development Evaluation DEval, Competence Centre for Evaluation Methodology, Bonn, Germany*

In light of a growing number of evaluations from various fields and disciplines, the potential for aggregating evidence for systematic learning and accountability beyond single interventions, projects or programs, increases. To date, the two most prominent approaches for the systematic aggregation of knowledge, meta-analysis and systematic reviews, are strong in terms of their systematic procedure but follow protocols that are only compatible with quantitative studies based on common methods, outcome indicators and explanatory variables. In contrast to this, there is an large body of evidence, which does not fulfill the strict requirements by systematic synthesis and is hence not considered for analysis. The authors argue that the development of more open but still rigorous systematic approaches of evaluation synthesis is key to avoiding blind spots in the world's evidence maps, and for enhancing the capacities for learning and accountability at an aggregated level.

In general, alternative approaches of evaluation synthesis offer ways for aggregating evidence obtained by different methods. However, more open synthesis designs challenge how to address the heterogeneous quality of single observations in a systematic manner. Meta-evaluations, which identify high quality studies, therefore often precede the synthesis of findings. While this procedure is rigorous in the sense that it aims to base syntheses on sound evidence only, it may come at the cost of arbitrariness, inadequate representation of evaluations in the synthesis of findings, and the loss of information and representativeness.

The authors propose an alternative design of quantitative evaluation synthesis. In particular, the authors suggest integrating findings from meta-evaluations as analytical weights in the evaluation synthesis. This technique overcomes drawbacks of cut-off point sensitive in- or out-approaches by explicitly taking into account the continuous distribution of methodological quality of a given base. The modified synthesis design is illustrated using data of a quantitative evaluation synthesis of sustainability assessment in the German development cooperation. Evaluation has to put more efforts in the development of suitable designs that allow to data aggregation and analysis as a basis for evidence-based decision making. The alternative approach of synthesizing quantitative findings from sources of varying methodological quality is a contribution to this overarching challenge.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 006 Mixing Methods in Systematic Reviews****O 017 - The Challenges of Screening and Synthesising Qualitative Research in a Mixed-Methods Systematic Review on the Impact of Agricultural Certification Schemes***D. Skalidou<sup>1</sup>, C. Oya<sup>2</sup>*<sup>1</sup> *University of East Anglia, School of International Development, Norwich, United Kingdom*<sup>2</sup> *SOAS University of London, Department of Development Studies, London, United Kingdom*

The number of mixed-methods systematic reviews has been growing in recent years in international development, following a general trend across different disciplines. By recognising the value of qualitative research in providing valuable evidence on causal mechanisms, barriers, facilitators and the importance of context, mixed-methods systematic reviews go beyond the 'what works' question of standard impact evaluation studies to also address questions on why, how and for whom interventions may or may not work. However, appropriate methods to screen and synthesise qualitative evidence in these reviews are still in a development phase, and the methodological literature dealing with reviewing qualitative evidence in the field of development studies is scarce and under-developed. This paper aims to contribute to this gap by discussing the methodological and practical challenges of including qualitative evidence in a mixed-methods systematic review in international development. In particular, this paper makes a contribution in terms of offering reviewers and users of systematic reviews a full account of the process of screening and synthesising evidence from a very large volume of heterogeneous qualitative studies. Using as an example a review on the effects of certification schemes for agricultural production, we report on each reviewing step, describing the problems encountered and solutions found. The paper proposes ways of extracting a large volume of data and integrating the qualitative synthesis with the evidence from the related quantitative effectiveness review.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 006 Mixing Methods in Systematic Reviews**

**O 018 - Knowledge Accumulation in Systematic Reviews:  
Advancing Strategies for Causal Generalization**

*S. Lemire<sup>1</sup>, C. Christie<sup>1</sup>*

<sup>1</sup> *University of California- Los Angeles, Department of Education and Information Studies, Los Angeles, USA*

The ability to generalize with confidence is a common concern in evaluation. The focus of the present paper is on the promise for and underlying nature of causal generalization in systematic reviews. Informed by the work of Thomas D. Cook, among others, the paper first distinguishes between two types of causal generalization: statistical and analytical generalization. Informed and illustrated by a recent meta-analysis, the paper then reflects on several fundamental issues related to statistical generalization – the most commonly used generalization strategy in systematic reviews. Advancing towards a broadening of our practice, the paper presents and considers four alternative strategies for causal generalization in systematic reviews: the quasi-sampling strategy (Cook, 1993), the realist strategy (Pawson, 2006), the EPPI-Centre strategy (Oliver et al., 2005), and the meta-modeling strategy (Lemire & Christie, 2018). In conclusion, the paper reflects on future steps towards promoting causal generalization in systematic reviews.

The proposed presentation emerges from and extends beyond a paper presented at the 2016 European Evaluation Society conference in Maastricht, a paper that was nominated for and later accepted for publication in *Evaluation* (Lemire & Christie, 2018). The present paper builds on and extends beyond Lemire & Christie (2018) by advancing a general framework for advancing our potential for causal generalization on the basis of systematic reviews. Moreover, and towards advancing our practice, the paper also illustrates and specifies how four different approaches (in markedly different ways) support causal generalization.

**References:**

Cook TD (1993) A quasi-sampling theory of the generalization of causal relationships. In: Sechrest LB and Scott AG (eds.) *Understanding Causes and Generalizing About Them*. San Francisco, CA: Jossey-Bass.

Lemire S and Christie CA (2018) Building the house of evaluation: Reconstructing the blueprints and bricks for knowledge building in evaluation. *Evaluation*, *in press*.

Oliver S, Harden A, Rees R, Shepherd J, Brunton G, Garcia J and Oakley A (2005) An emerging framework for including different types of evidence in systematic reviews for public policy. *Evaluation*, 11(4): 428–46.

Pawson R (2006) *Evidence-based Policy: A Realist Perspective*. London: Sage.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 007 Evaluating Innovation Policies and Programs****O 019 - Bridging The Gap Between Innovation Support and Socio-economic Impacts in Green Technology***V. Honoré<sup>1</sup>, V. Besrest<sup>1</sup>, P. Lefebvre<sup>2</sup>, S. Hooge<sup>3</sup>, A. Flichy<sup>1</sup>*<sup>1</sup> *Quadrant Conseil, Cooperative, Paris, France*<sup>2</sup> *Mines ParisTech, Competitivity, Paris, France*<sup>3</sup> *Mines Paristech, Scientific Management, Paris, France*

For more than a decade, massive support for green technology has arisen, associated with high expectations of socio-economic returns, finding its most well-known example in the Green New Deal launched by Obama in 2008. Consequently, there is a strong political demand to highlight the socio-economic effects of innovations policies, and public organizations in charge of such programmes are accountable for evaluating their impacts on economic activity and employment. To do so, a large panel of evaluation methods has been experimented. An analysis of the advantages and drawbacks of each one has led to the use of mixed methods (see for example the last evaluation of FP7 programme) adapted to the context of their implementation and the expected use of evaluation results.

But it is also increasingly acknowledged in the literature that, up until now, demonstrating evidence of significant impacts of innovation policies has proven to be very difficult, as evidenced by R&D collaborative projects in Europe (Dujardin, Lefebvre & al, 2017; Martin & al, 2011; Urraya & al, 2012). In addition, the counterfactual method, generally used for the attribution of such impacts, is limited to particular contexts and data availability and is not always adapted to produce knowledge about innovation paths with systemic and long-term results chain.

Consequently, which proposals can be made to better steer and estimate the socio-economic effects of an innovation support through theory-based methods? It is this question that the evaluation team of Quadrant Conseil and Mines Paristech has attempted to answer when developing an ex-post evaluation method for the Investissements d'Avenir programme operated by the French environmental and energy management agency (ADEME), based in particular on a literature review, an international benchmarking of impact evaluation practices in innovation agencies and a test on a sample of fifty projects supported by this programme.

Mobilizing concepts of innovation management (Osterwalder & Pigneur, 2010; Ries, 2011) to refine the intervention logic of the program enabled the evaluation method to pay more attention both to the intermediate results of the programme and to the necessity of preparing the commercialization phase from the beginning of the innovation process, alongside the technological development phase. This mixed method approach has enriched the evaluation of innovations projects, but still faces shortcomings when it comes to the attribution of long term impacts on economic activity and employment.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 007 Evaluating Innovation Policies and Programs**

**O 020 - Evaluation Methodology – Critical Foundations  
from Finland**

*J. Hyvarinen*<sup>1</sup>

<sup>1</sup> *Business Finland, Strategy and Impact Analysis, Helsinki, Finland*

In my paper, I will discuss several viewpoints, which should be critically considered before making evaluation analysis. Main point is how to interpret – in general – impact results in R&D and innovation policy. One important methodological issue, which should be raised is the results in risk-return portfolio. My paper presents how results differ when there is various risk-return setups in market failure context and how they differ from politician's dream world. Moreover, if take closer look to other methodological issues we find out several outcomes. First, goals of R&D and innovation projects differ. In other words, companies are heterogeneous and the consequences of these factors are not analyzed properly when considering proper control groups. Second, even if we are favor of econometrics, most of these models cannot estimate a risky behavior where standard deviation is high, i.e., results are shown as averages, not based on risk-return portfolio. Third, time lags are mostly too short and impacts cannot be measured properly because capabilities move forward to new sources in the economy. Finally, I present a framework how to use variety of methods for reaching useful impact results when considering analysis from multiple angles.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 007 Evaluating Innovation Policies and Programs

## O 021 - Human-Centered Co-Evaluation as a Lever for Learning and Evaluation Capacity Building (ECB)

*K. Hyytinen*<sup>1</sup>, *E. Saari*<sup>2</sup>

<sup>1</sup> VTT Technical Research Centre of Finland Ltd., Strategy and business intelligence, Espoo, Finland

<sup>2</sup> Finnish Institute of Occupational Health, Digitalisation, Helsinki, Finland

The traditional evaluation methods and measures are not able to capture neither the diversity of innovations in services and systems nor the multifaceted dimensions of performance resulting from these innovations. Reasoning is rooted in the “broad view on innovation” that highlights the interactivity, complexity and uncertainty of development and implementation of innovations. These arguments and perspectives affect both to the definition of innovations and to the evaluation of their effects and impacts (e.g. Kline & Rosenberg 1986; Lundvall 1992; Freeman 1991; Nelson & Rosenberg 1993; Dosi 1999).

Evaluation of innovations tends to focus on single values of technological progress and cost-efficiency. However, techno-economic perspective is too narrow to describe the multi-faceted nature of sustainable innovations (e.g. Djellal & Gallouj 2013). The paper outlines a human-centered evaluation framework, which captures the diversity of these innovations and their impacts. The developed framework bases on a multi-criteria framework which unfolds impacts of innovations or innovative experiments into six dimensions: *impacts on citizens, professionals and society as well as impacts on economy, integration of technology and services, and brand image*. Specific emphasis in the evaluation is on human and societal impacts, which are analysed symmetrically with the traditional techno-economic characteristics of innovation.

Traditionally, evaluations have been used as a feedback loop that measures the outputs of the innovation experiment or program in order to legitimize the inputs. However, if the sustainable change and learning from the experiment is prioritized, we need to involve all the value-creating actors; developers, users and potential distributors into a shared learning occasion. In the scientific evaluation discussion, there has emerged a concept of evaluation capacity building (ECB), to emphasize the importance of evaluation as an embedded element of organization’s learning practices or projects (Clinton 2014). Learning from the evaluations requires a process in which critical reflection on evaluation information (both qualitative and quantitative) takes place.

In human-centered co-evaluation method learning between different actors is enhanced by activity-theoretical classic principle of double stimulation (Vygotsky, 1978). The participants are offered a practical but inherently conceptual tool for analysing and making sense their shared activity and for making intentional decisions how to proceed in their “zone of proximal development”. The evaluation process between developers, users and potential actors, who may promote the innovation experiment provides an arena for learning and reflection along the development of the experiment. In the evaluation process the participants are offered a tool, which enables them to understand the service innovation in a wider context and long-term horizon. Reflexive tool use and collective evaluation has been previously used in developmental impact evaluation for innovation networks (Saari & Kallio, 2011). However, human-centered co-evaluation is based on a different learning method, which supports dialog between developers and potential distributors of the experiment. The dialog is supported by an aquarium method, which has been used in solving severe conflicts in a work community and also as an evaluation method. It is based on active listening: it instructs participants to listen, allows them to communicate and guides them to create further actions.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 008 Capturing the Impact of Cash Transfer Programs**

**O 022 - Overcoming Poverty and School Desertion. The Mixed Methods Evaluation Process of a “Programa Prospera” Pilot Program in Mexico**

*G. Sánchez Romero<sup>1</sup>, C. Mir<sup>2</sup>*

<sup>1</sup> ACEVAL, Executive Committee, Mexico City, Mexico

<sup>2</sup> Cocoa Services, Director, Mexico City, Mexico

This paper will present the methodological approach used in an evaluation of an innovative pilot program of Programa Prospera in Mexico. This comprehensive approach involved a qualitative and quantitative panel evaluation. Programa Prospera (henceforth Prospera) is the flagship public policy aimed at overcoming the intergenerational cycle of poverty through a conditional cash transfer model addressed to low income families in Mexico. In a nutshell, the pilot project consisted on redirecting the conditional transfer from the beneficiary mother to her sons and daughters in high school in order to pilot test whether these students would be better suited to take responsibility over their own school performance and avoid school desertion. The evaluation of this pilot was designed to understand how these processes perform in the household, and involved a mixed method approach based in qualitative inquiry and cost of opportunity study, analyzing at least two cohorts through time. In addition to presenting the methodological approach for the evaluation, this paper reports the main results, as well as the main challenges and areas of opportunity of these evaluation methodological approach.

The first section of this paper briefly introduces Prospera and its remarkable tradition of evidence-based decision making considering several evaluations on its twenty-year-old history. Then, it will present the main characteristics of the pilot program evaluated. On the next section, the paper will describe the methodological aspects of the design of the evaluation, highlighting its qualitative approach, based on a case study methodology. A case was composed by a beneficiary mother, her son or daughter, and close family members (father, siblings and other relatives living in the household). There were two data collection visits; one in 2016 and one in 2017, and the cases were mostly the same families visited in the two visits in order to control for the changes influenced by the pilot in one year of the intervention. The evaluation team visited four states in the north, center and south of the country, and elaborated twenty-nine cases evenly distributed throughout these regions. While the cases where the main source of analysis, the evaluation triangulated the data along with alternate sources, as focus groups and interviews with school staffers, teachers, bank clerks, and Prospera functionaries. The analysis answered nine evaluation questions, and the results of the first data collection in 2016 was contrasted with those of the 2017<sup>1</sup> collection.

The main results, combining the evidence provided by the cost effective study, was that there were none relevant changes provoked by the conditional cash transfer the students receive that would derive in reducing the school desertion of the cases studied, mainly because this transfer is not as high motivator as the moral and economic support they receive from their families if these believe education is important for their kids. The lessons learned from the evaluation will address the challenges of increasingly employing mixed methods in evaluations of pilot projects alike and other social development policies in the country.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 008 Capturing the Impact of Cash Transfer Programs****O 023 - Impact of Poverty Alleviation Through Unconditional Cash Transfers-Benazir Income Support Program in Pakistan**G. Mustafa<sup>1</sup><sup>1</sup> *Community of Evaluators Pakistan, Impact Evaluation, Islamabad, Pakistan*

Agenda 2030 envisages eradication of extreme poverty in all its forms everywhere. The Benazir Income Support Programme (BISP) is the main social assistance programme of Government of Pakistan and one of the largest in South Asia, serving 5.4 million beneficiaries. It was launched in 2008 as its flagship national social safety net initiative, in recognition that the existing instruments (Pakistan Bait-ul-Mal and Zakat) had limited coverage and were poorly targeted. For the Program, the grants were provided by the World Bank, DfID, ADB and USAID. However, about 87% of the program is being funded by Government of Pakistan out of its own resources. The BISP was launched with two main objectives: to cushion the adverse impact of the food, fuel and financial crisis on the poor; and a longer term objective of providing a minimum income support package to the poorest and to those most vulnerable to future shocks. The programme provides eligible families with unconditional cash transfers (UCT), originally set at a monthly value of PKR 1,000, raised to PKR 1,200 in July 2013, PKR 1,500 in July 2014 and PKR 1,566 in July 2015. The current rate since July 2016 is PKR 1612. The transfer is delivered quarterly, with the vast majority of beneficiaries receiving cash through the BISP Debit Card. By providing access to Computerised National Identity Cards (CNIC) and making BISP payments to the female head of beneficiary households the BISP made explicit the goal of the empowerment of women, which is complemented by the creation of BISP Beneficiary Committees (BBC) that provide a forum for beneficiaries. The programme established a National Socio-Economic Registry (NSER) through the use of an objective targeting system, with households targeted based on a Proxy Means Test (PMT) that attempts to provide an objective estimation of the level of income and welfare in all households in Pakistan and is summarised by the BISP poverty score. The NSER is now a database of more than 27 million households across Pakistan. The cash transfer is targeted at the poorest 25% of the population. Key intended impacts are: Increased consumption expenditure and poverty reduction; Women's empowerment; Increased household and child nutrition security; and Increased asset retention and accumulation. Issues, however, have been pointed out in the reports that the funds in some cases are not reaching the intended beneficiaries and are falling in the wrong hands. The issues of service delivery will be presented in the paper.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 008 Capturing the Impact of Cash Transfer Programs

## O 024 - Focusing on Double Vision: Are Proxy Means Tests Effective to Identify Future School Dropouts and the Poor?

*C. Crespo*<sup>1</sup>

<sup>1</sup> London School of Economics, Social Policy, London, United Kingdom

Conditional cash transfers (CCTs) have been targeted towards the poor. Thus, their targeting assessments check whether these schemes have been allocated to low-income households or individuals. However, CCTs have more than one goal and target group. Beyond poverty alleviation, CCTs seek to increase school enrolment. Hence, students at risk of dropping out of school are an additional target group.

This paper analyses whether one of the most common targeting mechanisms of CCTs, a proxy means test (PMT), can effectively identify the poor and future school dropouts. The PMT is compared with other approaches that use the outputs of a predictive model of school dropout. I built this model using machine learning algorithms (MLA) and rich administrative datasets from Chile.

The paper shows that using the outputs of the predictive model in conjunction with the PMT increases targeting effectiveness by identifying more students who are either poor or future dropouts. This joint targeting approach increases effectiveness in different scenarios except when social valuation of the two target groups largely differs. In these cases, the most likely optimal approach is to solely adopt the mechanism designed to find the highly valued group.

The paper provides novel contributions to the social policy targeting field. Overall, the results of the paper emphasise that targeting design, and assessments, must follow the goals of the policy and its consequential definition of target groups. Public officials that value the two described goals of CCTs equally may find opportunities for increased return on investment by modifying the targeting rules of these programs. Additionally, beyond providing one of the first machine learning applications of school dropout in a developing country, the paper shows that appropriate predictive models of this problematic are at hand for public officials. In contexts where countries are improving administrative records, this finding deserves attention.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 065 Designing Development For Change

## O 025 - Evaluators Unite: How RTI International has Leveraged Evaluation Expertise in Times of Critical Funding Constraints and Demand for Results

M. Chen<sup>1</sup>

<sup>1</sup> RTI International Research Triangle Institute International, Monitoring- Evaluation- Research- Learning- and Adapting Team- Global Health Division, Durham, USA

2018 has proven to be an uncertain year for evaluators working with U.S. Government funding, as we await to see how far and deep proposed cuts to the United States Agency for International Development (USAID) budget will be, the impending consolidation of the State Department and USAID, 200+ unfilled political appointees who are charged with many of the important decisions for our work, and other factors unbeknownst to us evaluators or even the rest of the world. The demand for results, accountability, and transparency are more critical than ever, and USAID implementing partners are more focused than ever to demonstrate their impact. These implementers are looking towards their data to see what's working and what's not and the gatekeepers of the data are those charged with overseeing and implementing the monitoring, evaluation, research, learning, and knowledge management of our programs as they serve as pillars of the results we produce. In a growing competitive market, where increasing numbers of new companies and nonprofits are entering each day and mergers and acquisitions overnight, how does an organization committed to research and evidence-based practice their place in this ever-changing market? They look to each other within the organization to find shared values, institutional knowledge, and support. The RTI International's Monitoring, Evaluation, Research, Learning, and Adapting (MERLA) Community of Practice (CoP) was created by three evaluator colleagues who found that they were dealing with similar challenges at work yet duplicating efforts. They decided to form a community of practice through internal funding and with senior leadership support. Communities of Practice (CoPs) can provide a theoretical grounding for a group of evaluators focused on understanding experience, increasing knowledge, and ultimately, improving and empowering our evaluation practice. They are designed to engage learners in a process of learning constructed around common interests, ideas, passions, and goals. In 2017, the RTI International's Monitoring, Evaluation, Research, Learning, and Adapting (MERLA) Community of Practice was created by three evaluator colleagues from diverse backgrounds – global health, international education, and governance and economic development – with the goal of creating a space where multiple experts working in M&E, research, training and capacity development, program management, operations, and other areas across our technical divisions could draw from existing expertise and experience to learn from each other. In our inaugural year, the RTI CoP has grown to 50+ members who have collaborated to conduct an internal survey of staff skills and expertise, design and deliver company-wide trainings such as "How to write Monitoring, Evaluation, and Learning (MEL) sections for proposals", and participated in the USAID Learning Lab's Collaborating, Learning, and Adapting (CLA) Challenge Week. Come join the co-founders as we talk about the RTI MERLA COP, share lessons learned and hear from others who are dealing with the same struggles of staying true to our mission to uphold transparency and accountability in a time of increased pressure, and how we plan on surviving these uncertain times by coming together to share resources, learn, and plan for the future.

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 089 Evaluating Educational Programs**

**O 026 - Learning for Adapting: Findings and Implications  
of a Peer Education Assessment in the Philippines**

R. Flueckiger<sup>1</sup>, E. Regan<sup>1</sup>, E. Dasmariñas<sup>2</sup>, G. Bede Benablo<sup>2</sup>, R. Benabaye<sup>2</sup>, C. Bisson<sup>3</sup>,  
J.O. Corciega<sup>2</sup>, C. Villa<sup>2</sup>, R. Colaco<sup>3</sup>, M. Chen<sup>3</sup>

<sup>1</sup> RTI International, International Development Group- Global Health Division, Research Triangle Park, USA

<sup>2</sup> RTI International, LuzonHealth Project, Pasig City, Philippines

<sup>3</sup> RTI International, International Development Group- Global Health Division, Washington- DC, USA

Evidence-based learning is critical for effective adaptation of project implementation. The use of appropriately designed operational research is key for assessing the successes and limitations of project implementation, that cannot be measure by planned monitoring and evaluation (M&E). Operational research provides defensible evidence for decision-making, strengthening the ability for projects to adapt. The provision of clear and statistically sound results serves as a catalyst for policy change at many levels.

According to the 2013 National Demographic and Health Survey, 10% of 15–19-year-old women in the Philippines are already mothers or pregnant with a first child. Teen pregnancy is most common among women of poor households, contributing to a cycle of poverty. The United Nations Population Fund determined that the Philippines is the only Asia-Pacific nation with a rising teen pregnancy rate. With a growing imbalance between working-age people and dependents there is concern that economic growth is limited in this country.

In 2014, the USAID supported LuzonHealth project in the Philippines designed and implemented school-based peer education sessions in 18 national high schools to promote adolescent health and provide information on teen pregnancy prevention. Trained peer educators facilitate sessions with their schoolmates that aim to provide correct, comprehensive and timely information on adolescent sexual and reproductive health.

To provide learnings for improving the project, we conducted a mixed method assessment. We found that students exposed to peer education demonstrated greater knowledge and awareness of pregnancy and sexual health, and more intention for safe behaviors than those not exposed. This evidence opens a channel for policy engagement to grow the program with Department of Education in collaboration with Department of Health, USAID and other stakeholders.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

## S 009 Toolkits for Evaluation Planning

**O 027 - Can Social Impact Evaluations be Standardized? Results from Introducing a Toolkit to Dementia Challenge Participants**G.Y. Reinhardt<sup>1</sup>, K. Chatsiou<sup>1</sup><sup>1</sup> University of Essex, Government, Colchester, United Kingdom

**Rationale:** In 1983, the UK Royal Institute of Public Administration identified at least three key areas for improvement in impact evaluation: coordination between governmental units; standardization across governmental units; and resourcing throughout the public sector.[1] In particular, the group determined a disproportionate need for these improvements in assessments of impact on difficult-to-measure concepts, such as resilience. Nearly 30 years later, researchers determined that few, if any, of these improvements had taken place.[2]

After the 2008 recession, the UK national government had endeavored to shift the responsibility of public service provision, including evaluation, from the national to the local level. New policies were to be based on high quality evaluations of previous policies and programs. Yet the shift illuminated a lack of experience in local-level public service impact evaluation,[3] and occurred without commensurate resourcing to help meet the new challenge.

**Objectives Sought:** In this paper we present pilot results of a new Program Evaluation Toolkit, which aims to address this shortfall by helping stakeholders design and begin an evaluation of a social program, assuming minimal expertise/knowledge.

The Program Evaluation Toolkit is a process for designing, planning, and coordinating impact and program evaluation that is useful to local authorities and voluntary organizations. A key advantage is that the toolkit helps an organization design an impact evaluation for a program intended to create impacts that are typically difficult to measure, such as improved wellbeing, reduced feelings of isolation, or increased empowerment, concepts that seem nebulous but are critical to building community and personal resilience to external shocks.

The Toolkit is delivered via web application and aims to:

- Help users design and plan public service delivery projects for local authorities; embed evaluation mechanisms into project design; create ongoing measurement mechanisms; construct data collection, evaluation & documentation protocols; broaden reach and impact beyond current locality;
- Initiate organizational culture change for program impact evaluation; Develop foundation for continued evidence-based policy design; disseminate toolkit;
- Raise awareness of evaluation as an integral part of program delivery, planning and administration; improve public service delivery by developing and disseminating the toolkit of templates and best practices.

**Brief Narrative and Justification:** A pilot Toolkit is available to 10 finalists for an Essex County Dementia Challenge Grant. The Toolkit platform enables these pilot users to design an evaluation of a prospective project to address dementia at the same time as they are crafting the application for funding the project. The Toolkit uses their input to generate customized evaluation plans, frameworks, questionnaires, data sets, information sharing protocols, and research ethics applications. Through surveys of the users and pre- and post-tests, we then assess whether and how views and capacities to design and conduct evaluations changes.

[1] Gray, Andrew, W. I. Jenkins, and Royal Institute of Public Administration. 1983. *Policy Analysis and Evaluation in British Government*. London: Royal Institute of Public Administration.

[2] Kell, Michael et al. 2013. *Evaluation in Government*. London: National Audit Office. Retrieved June 19, 2017 ([https://www.nao.org.uk/wp-content/uploads/2013/12/10331-001-Evaluation-in-government\\_NEW.pdf](https://www.nao.org.uk/wp-content/uploads/2013/12/10331-001-Evaluation-in-government_NEW.pdf)).

[3] Local Government Association. 2012. *Sector-Led Improvement in Local Government*. Retrieved June 19, 2017 (<https://www.local.gov.uk/our-support/our-improvement-offer/what-sector-led-improvement>).

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 009 Toolkits for Evaluation Planning**

**O 028 - Designing M&E in a Dynamic World: A Case of African Dynamic Environment**

*O. John Michael Maxel<sup>1</sup>*

<sup>1</sup> Uganda Evaluation Association, member, Kampala, Uganda

**Rationale:** Monitoring and evaluation undertakings come with challenges to organizations and evaluators in the process of selecting appropriate methods for assessments. Different organizations emphasize different methods others quantitative, qualitative or mixed methods. Emerging trends in Africa highlight challenges in the business environment. Deliberately picking up quantitative or qualitative or mixed methods without exploring contemporary conditions provides results that may mislead the evaluator. Emerging dynamics in the environments need evaluators to examine and understand the emerging challenges to come up with appropriate methodology. It will solicit promising approaches to overcoming limitations of current practices in turbulent times.

The choice of methods to be used for undertaking an evaluation depend on the situation on the ground their arguments that selection of methods depends on the philosophical foundation forms the basis for selecting the most appropriate research philosophy. An evaluator awareness of philosophical assumptions creates both innovation and creativity in the process of undertaking the inquiry. Whereas it is easy to recognize the methods, approaches, and techniques conducted by the researcher; questionnaires, interviews, focus group discussions but these are as a result of ontological and epistemological assumptions (Easterby et. Al., 2012). Quantitative, qualitative or mixed methods undertakings in evaluation have to be underpinned by the philosophical foundations of research.

**Purpose:** The purpose of this study is to provide the insight on appropriate mechanisms for the choice of evaluation designs and methods.

**Justification:** Monitoring and evaluation provide entities with a compass to guide programming, design, and execution of development in Africa. Monitoring provides insight towards improvement in the practice as the process used reveals strengths and weakness of the design, methods, systems, and implementation. This paper draws from the experience of the evaluator for more than a decade in evaluation government, University, and non-government organizations.

Global dynamics have substantially changed as a result of technology, politics, economics, socio-cultural dimensions continuously changing the choice of methods is critical in the establishment of the most appropriate approaches. The paper provides insight on the determination of appropriate methods for evaluation of contemporary phenomena. Whereas organizations emphasize qualitative, quantitative or mixed methods, the choice has to be underpinned by philosophical foundations. Monitoring and evaluation borrow actively from research, and the approaches in the process of selection of an appropriate method in the assessment in the dynamic global environment have to focus on the changing dynamics. Organizations or practitioners can't be casual in the selection of the methodologies for undertaking an evaluation. The paper prepared based on experience in evaluation in government, University and non-government programmes in Africa provides insight into the mechanisms for selection of appropriate M&E methodologies.

**Conclusion:** Emerging global trends in Africa requires rethinking of the choice of methodologies in the evaluation. Deliberately selecting methods without understanding the philosophical foundations can mislead well-intended evaluations. Emerging dynamics in the environments need evaluators to examine and understand the emerging challenges to come up with appropriate methodology. This is critical for evaluation and practice in Africa and beyond.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 009 Toolkits for Evaluation Planning

## O 029 - Towards a General Theory of Monitoring and Evaluation: The Measurement Of Organisational Capacity

*E.E. Goetsch*<sup>1</sup>

<sup>1</sup> University of the Witwatersrand, School of Law, Johannesburg, South Africa

**Rationale:** M&E's development as a scientific and professional discipline is hampered by the number of areas that remain hard to measure. One area that has been out of reach is the organisation's "capacity". Capacity is the key variable in the project's success. It is the first step in the production cycle (input->throughput->output->outcome) and the results-cycle (capacity->performance->delivery->impact). Practitioners want a robust method to monitor and evaluate capacity in order to direct funding, warn of gaps and predict failure before the project launches. This presentation fills that gap. It suggests a model of capacity and the algorithm for calculating it.

**Objectives sought:** To present a model of capacity that can credibly serve the profession as a standard tool for measuring the capacity of the organisation in 4 dimensions – its ability to effectively decide, design, deliver and document a project.

**Brief narrative:** The capacity model combines a number of disciplines. It compares the supply of labour with the demand for labour, where the supply is the sum of directors, managers and workers in position and the demand is the product of the number of governance units and key performance areas. The algorithm relies on a few constants; 3 being the size of a governance unit and 10 being the the number of key performance areas, making the demand for labour equal to 30. The effect is to render the organisational pyramid a rectangle. This method has dramatic results. The average capacity of World-Bank assisted projects in oil-rich Nigeria is 20%. This score is in line with the 30 implementation agencies evaluated in March 2018.

**Justification:** The model offers a quality standard for a universal measurement of capacity. It has many benefits. It helps evaluations help people to improve their lives and make our societies more resilient. It reduces unpredictability and complexity. It helps when designing and managing evaluations. It informs evaluation systems. It helps to rethink evaluation methods, design and criteria. It combines methods in evaluation. It addresses delivery risk in uncertain futures. It facilitates collecting and analysing data and reporting issues particularly in challenging contexts. It provides a dashboard that integrates ICT, M&E and managers. It enables evaluation to become foresight. It rewards flexibility and handles complexity. It develops the field of Evaluation to promote resilience and action in this critical time. It solves a challenge and offers opportunities for the evaluation field. It addresses some of the dilemmas and trends in professionalism, standards and ethical norms. It advances the theory and applies ethical values to evaluation. It promotes Evaluation Associations as custodians of a professional standard. It protects the independence of evaluators and our relevance and responsiveness. It supports the partnerships and stakeholders who make up the international development aid industry and greatly strengthens the communicating, using and embedding of evaluation. It is truly intersectional, since it connects M&E to every role-player in the organisation.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 010 Perspectives on Educational Evaluation**

**O 030 - Is Evaluating and Improving Leadership by Using Quality Standards and Indicators a “Hollow Bunny Trauma”**

*S. Čagran*<sup>1</sup>

<sup>1</sup> National School for Leadership in Education, Education, Ljubljana, Slovenia

Self-evaluation, supplemented by various forms of external evaluation and accountability systems, is one of the main tools for identifying and ensuring quality in education. The responsibility for the implementation of self-evaluation processes and the implementation of improvement measures remains in the domain of a particular school, since self-evaluation is defined as a process initiated and implemented by the school in order to describe and evaluate its own operation. In the cyclical process of self-evaluation schools define areas of assessment of their own quality, identify strong and weak points, set goals, prepare action measures for improvements and monitor the effects of improvements introduced.

Introducing leadership standards and indicators in the national context is based on the fact that school systems are searching for different ways on how to ensure that schools are run by great leaders. Emerging standards, that will be presented, are based on literature review, national legislative framework, examples of good practices of other educational systems and feedback of principals. The questions that however still remains is, how to present them, so that school leaders will actually use them and therefore improve their practice. For this reason a research, using focus groups of school leaders, was conducted to clarify the purpose and method of implementing the standards. Presented research findings will point out some starting points for a discussion on the potentials, strengths, challenges, dilemmas and threats of standards implementation for evaluating and improving school leadership.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 010 Perspectives on Educational Evaluation

## O 031 - The Impact of Education Programmes, Evidence from an Evaluation of the Integrated School Performance Improvement, Review and Engagement Project, Malawi

A. Mcpherson<sup>1</sup>, S. McDonald<sup>2</sup>

<sup>1</sup> Bentley University, The Jeanne & Dan Valente Center for Arts & Sciences, Waltham, USA

<sup>2</sup> Bentley University, Global Studies, Waltham, USA

The paper examines the effectiveness of key features of the 'Integrated School Performance Improvement, Review and Engagement' (INSPIRE) project in Malawi, drawing on the findings of a recent endline evaluation of the project. In doing so the paper considers whether the evidence is consistent with international evidence on the effectiveness of education programmes on children's learning and school participation.

In low and middle-income countries (L&MICs), improvements in children's school enrolment rates have slowed down considerably since 2004 with around 263 million children and youth still out of school. Access to schooling has also not translated into an improvement in children's learning outcomes in many L&MICs. In addition, UNESCO's 2014 Education for All global monitoring report comments that approximately 250 million children in L&MICs cannot read, write or do basic maths.

As a low-income country, the Malawi education system faces several challenges, including: inadequate school facilities, high dropout rates, high pupil-teacher ratios, low completion rates, and low learning achievement for children in poor rural areas (where HIV/AIDS prevalence is high).

The Scottish Government funded INSPIRE project sought to address challenges in the Malawi education sector and improve the performance of Malawian schools (and the impact of the Malawian Ministry of Education Science and Technology (MoEST) and District Education Offices (DEO)), through developing integrated planning, multi-stakeholder accountability and the provision of effective support. The project also aims to develop and demonstrate a consolidated district school monitoring and support system clarifying the rights and responsibilities of all stakeholders.

This paper investigates whether outcomes associated with the main features of the INSPIRE project are consistent with the international effectiveness evidence. The evaluation used of a combination of desk research and primary research methods, reflecting the study aims, the range of stakeholder groups, and the resources available. The main methods included: face-to-face, in-depth, semi-structured interviews with selected key stakeholders; focus group/s; case study research; and a questionnaire survey of selected officials and stakeholders.

There is preliminary evidence for a range of benefits being experienced by participant schools in the INSPIRE project. However, the consistency and nature of evidence for benefits is, provisionally, in some cases, at odds with the international effectiveness evidence. The paper highlights the limitations in making comparisons with international evidence due to differences in research focus, country context, and evaluation methodology. In particular, the paper highlights challenges in establishing counterfactual evidence in small scale projects and in making comparisons with wider impact evaluation evidence on the effectiveness of education programmes on children's learning and school participation.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 010 Perspectives on Educational Evaluation

## O 032 - School Self-evaluation Capacity Building – Should We Model it?

*M. Brejc<sup>1</sup>, S. Čagran<sup>1</sup>*

<sup>1</sup> National School for Leadership in Education, National School for Leadership in Education, Ljubljana, Slovenia

Self-evaluation in schools is not a matter of question – it is a fact! And by recognizing the value of self-evaluation the question of capacity building becomes an issue. Capacity building is necessary for quality and sustainable use of self-evaluation in schools with regards to teachers and school leaders not being professional evaluators, meaning that knowledge, skills and attitudes are not self-evident.

Literature and research review referring to a wider concept of evaluation capacity building (ECB) defines its general aim as sustainable practice of quality evaluation. The concept and practice of ECB are still in the 'phase' of inquiry, there is little literature or research on school self-evaluation capacity building.

Based on wide literature review of ECB and research on school self-evaluation training as one of ECB strategies school self-evaluation capacity building (sECB) model (Brejc 2014) will be presented and discussed. sECB is defined as a) the process of gaining and/or improving knowledge, skills and attitudes of teachers and leaders for self-evaluation, b) establishment of institutional frameworks and structures for conducting effective and sustainable self-evaluation aiming at improving the quality of school performance, and c) framing of the system demand, guidelines and support for school self-evaluation efforts. As such it addresses individual (eg. teacher, school leader), organizational (eg. school) as well as system level emphasising the need to link them. Model will be presented in the framework of evaluation culture, school improvement and (professional) accountability. Strengths, weaknesses, challenges and dilemmas will be discussed from a point of view of specific national school system.

Brejc, M. 2014. Krepitev zmožnosti za samoevalvacijo v šolah (Building capacity for school self-evaluation). Kranj: Šola za ravnatelje.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 010 Perspectives on Educational Evaluation**

**O 033 - Reconsidering Interventions as Time-Varying Events:  
Lessons Learned from an Evaluation of Early Grade Retention**

*S. Lehouelleur*<sup>1</sup>

<sup>1</sup> Ghent University, Department of Applied Mathematics- Computer Science and Statistics, Bruxelles, Belgium

Impact evaluations often consider the treatment intervention of interest to be a single event (in time). These evaluations ignore that the effect of the treatment intervention might develop and be influenced by repeated intervention events over time and in effect be better viewed as a time-varying treatment, that is, a causal path from intervention onset – through intermediate intervention events and outcomes – to long-term impact.

Based on a recent longitudinal evaluation of early grade retention, this presentation illustrates the benefits of considering early grade retention as a time-varying treatment intervention, whereby intermediate treatment confounders are adequately controlled for. The presentation first considers the conceptualization of early grade retention as either a single treatment intervention or a time-varying treatment intervention. Informed by this distinction, the presentation then compares and contrasts impact estimates based on each conceptualization. Finally, the presentation concludes by considering the implications of the findings for impact evaluation.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 050 Evaluating Livelihood and Community Programs**

**O 034 - Adoption and Impacts of Improved Beekeeping Technologies in the Miombo Woodland of Tanzania**

*N. Kuboja<sup>1</sup>, A. Isinika<sup>2</sup>, F. Kilima<sup>3</sup>*

<sup>1</sup> Ministry of Agriculture, Research And Development, Dar es Salaam, United Republic of Tanzania

<sup>2</sup> Sokoine University of Agriculture, Institute of Continuing Education, Morogoro, United Republic of Tanzania

<sup>3</sup> School of Agricultural Economics and Business Studies- Sokoine University of Agriculture, Agricultural Economics and Agribusiness, Morogoro, United Republic of Tanzania

This paper analyzes the adoption and impacts of improved beekeeping hives on household's income among beekeepers in western Tanzania using cross-sectional data obtained from a sample of 198 beekeeper households. Using propensity score matching and endogenous switching regression models, the paper shows that adoption of improved hives had a significant gain in household income accrued from beekeeping. An analysis of the determinants of adoption revealed age of the household head, years of formal schooling, access to credits, access to extension services, training and experience in beekeeping as key factors influencing adoption of improved hives. This study supports the need for using improved beekeeping technologies for improved productivity as well as household income among small scale beekeepers. Thus, efforts to improve access and use of improved beekeeping technologies should be part and parcel of income poverty reduction strategies among rural population in the study area where many households are depending mainly on beekeeping. The paper concludes that policies that enhance diffusion and adoption of improved hives should be central to income poverty reduction strategies in Tanzania.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 012 Training and Professionalisation of Evaluators**

**O 035 - Evaluator Competencies and University-Based Evaluation Education In 2018: Global Perspectives and Implications for Professionalism**

*J. Lavelle*<sup>1</sup>

<sup>1</sup> *University of Minnesota-Twin Cities, Organizational Leadership- Policy- & Development: Evaluation Studies, Minneapolis, USA*

The field of evaluation has grown over the years. This has been hallmarked by a rise in the number of evaluation societies across the world, greater demand for evaluation services in a variety of contexts, and growth in the number of universities awarding degree concentrations (e.g., certificate, master's, or doctorate) in evaluation (LaVelle, 2014). Questions persist, however, about the alignment between the skill set desired by the international evaluation job market, the competencies being developed through formal courses of study, and the possible implications for the profession if or where there is misalignment.

Concurrently, the pre-service preparation of evaluators through university-based evaluation education programs has been the subject of regular inquiry for the professional evaluation associations. This interest often culminates in the publication and dissemination of a program directory. Although the profession of evaluation has evolved greatly, the last published directory of programs in the United States was in 2010 by LaVelle & Donaldson, and European programs in 2006 by Bewyl & Harich. However, with the renewed proliferation of evaluation societies across the world (Donaldson, 2006; Donaldson, 2015) comes a reinvigorated interest in the pre-service preparation of evaluators.

Building on an analysis of university-based evaluation programs across the world and previous work by LaVelle (2014) and LaVelle & Donaldson (2010; 2015), this presentation will share the results of an ongoing mixed-methods study contrasting the desired evaluation skills in an international market with the skills being developed in degree-granting university programs. It is anticipated that there will be areas of strong alignment as well as opportunities for better alignment. Implications for the professionalization of evaluation will be discussed.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 012 Training and Professionalisation of Evaluators

## O 036 - Training Evaluators to be Resilient and Action Oriented: Contrasting Perspectives on Teaching

*C. Lovato*<sup>1</sup>, *D. Pratt*<sup>2</sup>, *S. Rusticus*<sup>3</sup>

<sup>1</sup> University of British Columbia, School of Population & Public Health, Vancouver, Canada

<sup>2</sup> University of British Columbia, Centre for Health Education Scholarship, Vancouver, Canada

<sup>3</sup> Kwantlen Polytechnic University, Psychology, Vancouver, Canada

**Rationale:** In evaluation, if we are to reconsider the theories, practices and fundamental questions regarding our role in fostering change, we must reconsider the teaching of evaluation. How can evaluation be taught so that new ideas about evaluation are carried forward into the next generation? This presentation will make the case that some orientations to teaching are more likely to replicate the past than reimagining a future.

#### **Objectives:**

1. Consider differences in the way teachers of evaluation perceive their role and responsibilities to trainees
2. Explore perspectives on teaching evaluation that support resilience and an orientation to action.

**Narrative/Justification:** Are some pedagogical practices for educating evaluation trainees better suited to helping people improve their lives and make our societies more resilient? Are some pedagogical practices more likely to replicate the past than imagine the future? To address these questions, an understanding of the beliefs and intentions of those who teach evaluation is needed. Research shows differences in the ways teachers perceive their roles and responsibilities. These differences have been demonstrated within and across a wide range of disciplines. This presentation will explore perspectives on teaching that are well-suited to rethinking evaluation in ways that respond to the turbulence of our times.

Pratt's Five Perspectives on Teaching (Pratt, Smulders & Associates 2016) will provide the means for examining these questions. According to Pratt et al., there are five qualitatively different teaching perspectives. Each perspective is based upon an interrelated set of intentions and beliefs that give direction and justification to a teacher's actions. Within the **transmission** perspective a teacher's primary responsibility and commitment is to present content accurately and efficiently. **Apprenticeship** teachers are experienced practitioners of what they teach; they socialize learners into an existing community of practice. **Developmental** teachers' primary commitment is to develop and foster the growth of complex and sophisticated forms of reasoning. **Nurturing** teachers build a trusting relationship between teacher and learner in which there is a balance of intellectual challenge and emotional support. While a primary commitment of a **Social Reform** perspective is to critique and change the status quo. Social Reform teachers expect students to question existing practices and encourage them to take action to improve their own lives and those of others. We propose that the Developmental and Social Reform perspectives are best suited for the current era of challenging times; and Transmission and Apprenticeship are more likely to promote existing theories and practices.

We are using a validated instrument, the *Teaching Perspectives Inventory* (Collins & Pratt, 2011) to profile the orientations of those who teach evaluation as we look for patterns that may fit the call for dealing with turbulent times better than others. Descriptive findings, profiling the perspectives of teachers that train future evaluators to be resilient and action-oriented practitioners will be presented in contrast to those teachers that are more likely to replicate past practices.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 012 Training and Professionalisation of Evaluators

## O 037 - Professional Networks and Professionalization in the Evaluation of Public Policies in Ibero-America

*E. Santiago*<sup>1</sup>

<sup>1</sup> 1 Red Internacional de Evaluación de Políticas Públicas RIEPP. 2 Auditoría General de la Ciudad de Buenos Aires. AGCBA, Dirección General de Control de Cultura., Buenos Aires, Argentina

The evaluation of public policies, although it is an activity valued in modern public administration, is the first victim of budget cuts in times of crisis such as those we are currently going through. Both in times of boom and retraction, networks of evaluation professionals have a very valuable role to maintain its validity and relevance in the public agenda.

These organizations, networks and societies grouped at the national level or by other identity attributes, are central drivers of the culture and practice of evaluation; and the contributions that this activity can offer to the policies.

This paper offers a space for reflection on the contributions of the networks of evaluators.

Its purpose is to share progress and achievements in its efforts to promote the culture of evaluation of public policies; as well as the difficulties and pitfalls that its activity faces.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 013 Breaking the Boundaries of Evaluation

## O 038 - Crossing Evaluation Boundaries in the Face of Uncertainties: A Taxonomy to Commission, Conduct and Use Better Quality and Resilient Evaluations

*M. Tarsilla*<sup>1</sup>

<sup>1</sup> UNICEF- West and Central Africa Regional Office WCARO, Regional Evaluation Advisor, Arona NO, Italy

Regardless of the specific role that we may play in the course of an evaluation, personal and technical boundaries affect our behaviors and decisions throughout the process (e.g. evaluation commissioners might privilege certain criteria when developing Terms of Reference; evaluation practitioners might promote the use of one methodology over and over again based on their specific academic background).

However, what happens when the contexts – in which evaluations get commissioned, conducted and hopefully used- become more and more unpredictable? How could evaluation prosper at a time of uncertainty and how could evaluation practitioners – both the more seasoned and emerging ones- become more resilient? It is the presenter's conviction that "crossing evaluation boundaries" (that is, going beyond the linearity and rigidity of certain evaluation practices and mental models that characterize our practice) is needed in order to preserve the relevance and credibility of the evaluation profession in the future. That a critical discussion on this theme ought to take place is also confirmed by the paucity of literature available on the topic (a cursory online search of the term "crossing boundary in evaluation" in march 2018 only yielded one result). In an effort to address this gap, the presentation will engage the audience in a critical reflection on how to rethink and overcome some of the methodological, conceptual and paradigmatic boundaries underlying their own evaluation practice. In doing so, the session will not only enrich the conversation on this year's conference theme but also foster new thinking on how to promote more resilience-oriented evaluation in the future, both within and outside of the European Evaluation Community.

Drawing upon a systematic review of literature of evaluation practices in volatile and unpredictable contexts, the presenter will develop a taxonomy of evaluation practices and methodological choices that go beyond conventional conceptual and methodological boundaries. Such rather comprehensive and inclusive taxonomy will build on an earlier draft developed by the presenter in 2017 and will benefit from the contribution of a large number of evaluation practitioners, managers, commissioners and users, who will be asked to comment on it through a global survey and a series of follow-up online discussions held on a variety of evaluation listserves during the months preceding the conference.

Each one of the categories discussed in the taxonomy will feature some real-world examples of "crossing boundaries" in evaluation in several regions of the world and will include a list of concrete recommendations that the audience may follow in order to cross their own personal and professional boundaries in evaluation so as to make their practice more resilient.

Furthermore, given the presenter's personal and professional engagement on issues pertaining to gender equality, equity, intersectionality and human rights in evaluation, the presentation will provide sufficient details on how "crossing boundaries" in evaluation could contribute to more just and respectful evaluation practices, especially in light of the existing gender imbalances and power asymmetries observed in past evaluations.

Participants' comments shared via Twitter during the presentation will also be discussed towards the end of the session.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 013 Breaking the Boundaries of Evaluation

## O 039 - Revisiting the Lessons of the Global Decolonisation Movement for More Effective Research and Evaluation

*F. Cloete*<sup>1</sup>, *C. Auriacombe*<sup>2</sup>

<sup>1</sup> University of Johannesburg, Public Governance, Johannesburg, South Africa

<sup>2</sup> University of Johannesburg, Public Management and Governance, Johannesburg, South Africa

**Rationale:** There is increasing global resistance in many circles against a perceived Euro-centric value hegemony in knowledge generation, implementation and evaluation in many developing countries. This situation is perceived to impose outdated and inappropriate policies on these countries and need to change to more appropriate processes and results to improve conditions in those countries in the 21<sup>st</sup> century. 21<sup>st</sup> century societies are generally acknowledged to be increasingly complex and volatile due to increases in inter alia the heterogeneity of populations, demands for change, and faster communications technologies. This occurs in parallel with a growing decrease in resources, time and therefore capacities for meeting the above challenges in the most effective ways.

**Objectives:** The envisaged focus of the paper is to learn some lessons from the impact of historical colonial value systems and practices in current knowledge generation, transfer and application processes and results in Africa (especially in South Africa). The objective is to identify concrete directions towards 'decolonising' research and evaluation processes and products to be more relevant, appropriate and therefore more effective to address the above 21<sup>st</sup> century challenges better, not only in lesser developed countries but also in traditionally more developed western nations. The paper address the following issues: 1) How is the need for decolonisation conceptualised?, 2) What aspects of prevailing research and evaluation-related dimensions need to be 'decolonised'?, and 3) What lessons can be learnt from the prevailing debates on these issues in the African context for improved research and evaluation outcomes in general.

**Findings and Conclusions:** The findings of the paper include that watertight distinctions between Euro-centric and non-Euro-centric (especially Africa-centric) approaches do not always exist, and that it is risky to over-generalise such alleged distinctions in many cases. However, there are historical and current research and evaluation processes and practices that are inappropriately applied in African and other non-Western societies and in need of change. A single change recipe does not exist, and different solutions should be considered in different situations in order to achieve more effective and resilient research and evaluation impacts on turbulent, complex contemporary societies. The conclusions of the paper include that the current decolonisation discourse is largely rhetorical, outdated, negative and ineffective. A more relevant, re-focussed, positive, pragmatic, resilient and integrative approach is required to improve the potential impacts of research and evaluation on societal change. The development of transformative, trans-disciplinary, developmental, culturally and context specific and sensitive, mixed research and evaluation approaches, designs and methods, are emerging good practices in the right direction. These trends can and should be expanded and applied in more consistent ways to move beyond the empty negative rhetoric of the current decolonisation discourses to a more constructive and empowerment-focussed outcome. The paper identifies the main concrete strategic directions and strategies needed to achieve such envisaged outcomes. It is in the public interest, relevant to EES members, could improve creativity, innovation, skills, gender equality and resilient societies and is hopefully of adequate professional quality.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 014 Evaluating for Resilience and Climate Change

## O 041 - Evaluators and VOPEs With Resilient Cities: the ABC of What to Know and Do from Day 0

*P. Smits*<sup>1</sup>

<sup>1</sup> U Laval, Management, Canada

**Rationale:** Resilient cities are now given high visibility. Preparing to face contingencies for municipalities calls for quite a complex intervention. Municipalities need to better coordinate across sectors, eventually with the integration of provincial and national responsibilities under a municipal leadership, etc. Municipalities have to put in place structures and procedures to inform and train citizen coming from heterogeneous background, knowledge and sensibility. Municipalities also need to plan and anticipate unforeseen events, planning upon the unknown. Overall, municipalities have to pool considerable efforts to take on unknown paths.

The contribution of evaluators can be technical when measuring the impacts of resilient cities. The role can also extend to the resilient cities' strategy development. Evaluators can be part of the earliest thinkers and the positive energy of emerging groups ready to brainstorm, support and launch resilient cities.

This presentation explores how the evaluation community can be part of the initiators of resilient cities, bringing evaluation to the earliest steps of decision making.

#### Objectives sought:

- Argue for the benefits to reverse translate in the field of evaluation (translating from municipal concerns for governance of resilient cities to the field of evaluation, instead of bringing evaluation in the world of municipal decision makers.)
- Provide VOPEs, evaluation practitioners and managers with indicators to follow early stages to initiate resilient cities
- Provide VOPEs, evaluation practitioners and managers with adjusted evaluation processes to follow early stages to initiate resilient cities
- Know examples on how to translate municipal stakeholders' concerns into evaluation language related to resilient cities' evaluation criteria, resilient cities contextualized evaluation methods, evaluation processes related to evaluating gender aspects, young, sustainable development goals in resilient cities.
- Reflect as a community of evaluators upon our complementary capacities to reach out to non evaluation users at municipal level

**Brief narrative and justification:** We draw from two sources of data: a focus group with 30 individuals interested in launching a resilient city in Montreal Canada and how to measure accomplishments from the earliest steps, and a discussion between 4 VOPEs in evaluation interested in supporting the emergence of a resilient city in central Asia.

We will present : 1. Resilient cities-informed indicators, and 2. Resilient cities-informed evaluation processes.

We will conclude on the role of evaluators in the early stages to build resilient cities. The roles are two-fold: evaluation entrepreneur or how the evaluation community can advertise and advocate for evaluation in the context of resilient cities, evaluation broker or how the evaluation community can bridge municipal stakeholders concerns with evaluation tools and processes. We will also discuss on the role of VOPEs in the early stages to build resilient cities: VOPEs preaching to non users of evaluation in municipalities; VOPEs peer-to-peer networking on best practices and capacities needed by VOPEs to reach out to resilient cities developers.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 014 Evaluating for Resilience and Climate Change**

**O 042 - Using Theory of Change Models to Promote Inclusive Climate Action in Cities – the Urban Climate Action Impact Framework (UCAIF)**

*K. Attstrom*<sup>1</sup>

<sup>1</sup> Ramboll Management Consulting, European Policy, Saint Nazaire de Pezan, France

The Urban Climate Action Impacts Framework, UCAIF, is intended as a user guide with common principles, taxonomy and guidelines for approaching the mapping and assessment of wider impacts of climate actions in cities. The work aims to catalyse medium-term development of the tools, resources and evidence cities need to make the case for climate action by linking climate actions to the other priorities they face, such as health, poverty reduction, economic development and equity, including links to the SDGs. The work was done in collaboration between Ramboll and the C40 Cities Climate Leadership Group, and was based on literature review and consultations with institutional stakeholders working in the field of climate action, inclusion and equity.

Above and beyond being of direct relevance to meet climate challenges in cities, the work undertaken serves to illustrate how evaluation “theory of change” models can be used to serve broader purposes than gathering evaluative knowledge. The paper will also make the case for using logic models to promote systems thinking, by taking into account co-benefits and negative impacts from a wider perspective.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 014 Evaluating for Resilience and Climate Change**

**O 043 - Tracking Contextualised Adaptation Across Scales:  
The TAMD Framework**

*E. Beauchamp*<sup>1</sup>

<sup>1</sup> *International Institute for Environment and Development, MEL / Climate Change, London, United Kingdom*

Focus on local-level climate adaptation investments have increased in recent years, bringing the need to evaluate how resilience has been affected across the different scales through which climate financing is channelled. For resilience interventions to be sustained, adaptation needs to occur at multiple levels of a system. Evaluating such interventions and policies is particularly challenging given that for adaptation to be done well, it must be tailored to changing contexts – rendering top-down indicators inappropriate. Yet current results frameworks on resilience rarely used locally defined outcome at more than one level, and often hinder the generation of learning to promote institutional and local adaptation.

In this presentation, I introduce the Tracking Adaptation and Monitoring Development (TAMD) conceptual framework. Developed by the International Institute of Environment and Development (IIED), TAMD can be used in many contexts and at many scales to assess and compare the effectiveness of interventions that directly or indirectly assist populations in adapting to climate change. It combines top-down institutional indicators with bottom-up vulnerability indicators, acknowledging and addressing problems associated with changing climatic baselines. It provides an explicit framework for two paths, or “tracks”; Track 1 entails assessing the capacity of institutions to undertake effective climate risk management actions (also called top-down), and Track 2 entails assessing impacts of interventions aimed at reducing vulnerability and the extent to which such interventions keep development on track (development performance or bottom-up).

TAMD can be used by countries to evaluate how far, and how well, climate risks are managed at international, national and sub-national scales. It uses vulnerability and development indicators to assess whether development outcomes bring better local climate resilience, and whether that aggregates at larger scales to contribute to climate-resilient development. The TAMD framework concentrates on developing robust and bespoke frameworks tailored to national circumstances by partnering with government agencies responsible for delivering social and economic development, using existing information on development progress, introducing new thinking and frameworks for assessing climate risk management, and fostering shared learning and a community of practice among public sector staff across developing countries. TAMD can be used by government officials at local and national levels, NGOs and development partners, fostering capacity-building for M&E of adaptation needed to create more resilient, climate-sensitive institutions.

Along with the framework, we present results from applying TAMD over three years to evaluate effects of the Decentralized Climate Finance project in Mali and Senegal, from 2015 to 2017. We discuss practical implementation steps of the TAMD framework, challenges and implications for long-term data collection for interventions aiming at improving resilience to climate change.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 014 Evaluating for Resilience and Climate Change**

**O 044 - Opportunities and Challenges for Using Performance Measurement Information in Support of Resilience: The Case of Local Governments in Greece**

*P. de Lancer Julnes<sup>1</sup>, S. Plimakis<sup>2</sup>, E. Serre<sup>3</sup>*

<sup>1</sup> Penn State Harrisburg, Director- School of Public Affairs, Middletown, USA

<sup>2</sup> University of Peloponnese, Political Science, Corinth, Greece

<sup>3</sup> Penn State Harrisburg, Public Administration, Middletown, USA

The turbulence of current times is prompting calls for replacing the concept of sustainability with the concept of resilience. They argue that sustainability traditionally has meant practices and solutions expected to lead to conditions of stability and reversal of past mistakes (Ahern, 2011). Turbulence, however, calls for resilience – defined as recovery and transformation (Shaw, 2012). Resilience is the ability to adapt to the unexpected.

In the context of local government resilience, the Rockefeller Foundation 100 Resilient Cities initiative describes city resilience as “the capacity of cities to function, so that people living and working in cities – particularly the poor and vulnerable – survive and thrive no matter what stresses or shocks they encounter “(100 Resilient Cities, 2018, n.p.). These stresses and shocks could be man-made or natural events. Research suggests that some of the more broadly applicable characteristics that prevent municipalities from being resilient include rigidity, informality, decentralization, traditionality, and fragility. According to Pelling and Manuel-Navarrete (2011), existing governance regimes tend toward actions that support the status quo and stifle flexibility, impeding the ability for regimes to create sustainable development.

One key element of resilience strategy is having the capacity for making decisions. This capacity for decision-making requires having information about the achievement of goals and strategies and using the information. For example, in its resilience strategy developed as part of the Rockefeller Resilient Cities initiative, the municipality of Thessaloniki identified the development of a performance-based management system (PBM) as one of its objectives to achieve the goal of building “a dynamic urban economy and responsive city through effective and network governance” (p. 90). In doing this, the resilience team sought to develop mission, goals, objectives and performance baseline that can help the municipality improve the quality and delivery of city services and support priority sectors. As suggested above, a PBM system requires performance measurement – the regular collection and analysis of data on a program or service inputs, outputs, and outcomes.

This paper will explore the use of performance measurement information by Greek local government employees and officials. Using data from a survey, we will discuss the challenges and opportunities that local governments in Greece currently face with regard to performance measurement and the implications on their ability to build resilience.

**References:**

100 Resilient Cities (2018). Resources. (<http://www.100resilientcities.org/resources/#section-1>)

Ahern, J. (2011). From fail-safe to safe-to-fail: Sustainability and resilience in the new urban world. *Landscape and Urban Planning*, 100(4), 341–343. doi:10.1016/j.landurbplan.2011.02.021

Pelling, M. & Manuel-Navarrete, D. (2011). From Resilience to Transformation: the Adaptive Cycle in Two Mexican Urban Centers. *Ecology and Society*, 16(2).

Resilient Thessaloniki. (2018). [https://www.100resilientcities.org/wp-content/uploads/2017/07/Thessaloniki\\_Resilience\\_Strategy\\_PDF.pdf](https://www.100resilientcities.org/wp-content/uploads/2017/07/Thessaloniki_Resilience_Strategy_PDF.pdf)

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 015 Applying Contribution Analysis in International Development****O 045 - A Practical Approach to Address the Attribution Challenge in Complex Interventions: Evaluating Impact of a Market Systems Development Programme**G. Ton<sup>1</sup>, A. Koleros<sup>2</sup><sup>1</sup> *Institute of Development Studies, Centre for Development Impact, Brighton, United Kingdom*<sup>2</sup> *Palladium, Research- Monitoring and Evaluation, London, United Kingdom*

This paper presents the methodological design and mid-term results of a theory-based mixed-methods approach to assess the impact of the Private Enterprise Programme Ethiopia (PEPE), an adaptive programme that addresses key constraints in the leather and garment industries in Ethiopia through a market systems development approach. We use Contribution Analysis as the overarching framework to verify the main interlinked causal assumptions in PEPE's theory of change. The evaluation employs a number of innovative methods to achieve this. Nested in the mixed-method design, we use Process Tracing as a structured method of critical inquiry in the qualitative case studies to verify whether the intervention is indeed a non-redundant contributory factor to the improved service provisioning. In the firm survey we apply survey modules to compute so-called Contribution Scores on various intermediate outcomes (firms practices). The CGE modelling extrapolates the within-sample effects to effects in the wider economy. This method helps to match the need of the commissioner (DFID) to report on harmonised indicators of development impact with the real-world limits to the span of direct influence of the development support (changed practices in service providers). We will explain how we analysed the survey data to define the scenarios that are used for modelling these higher-level development impacts. We conclude with a discussion on strengths and weaknesses of this approach to assess the impact of adaptive programmes in complex systems.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 015 Applying Contribution Analysis in International Development****O 047 - Contribution Analysis and Translation Theory to Measure Use and Influence of a Evaluation: Rethinking Evaluation Utility Standards**D.M. Franco de Abreu<sup>1</sup>, A. Brouselle<sup>2</sup>, E. Moreira dos Santos<sup>1</sup>, G. Cardoso<sup>1</sup>, E. Artmann<sup>1</sup><sup>1</sup> Fundação Oswaldo Cruz, Escola Nacional de Saúde Pública, Rio de Janeiro, Brazil<sup>2</sup> University of Victoria, School of Public Administration, Victoria, Canada

The restriction of resources for social policies in Brazil (JANUZZI, 2001), have increasingly demanded of evaluations the commitment to their applicability and effectiveness for change. Even though there was an enhance on health evaluation in Brazil during the last years, one of its challenges is the incorporation of findings from these evaluations, specially to improve interventions and to fit timely decision making. This paper aims to present the metaevaluation findings from a tuberculosis control program's performance evaluation implemented in three Brazilian cities. The metaevaluation problematize the utility and influence standards. Both, have being concerning evaluators and also addresses evaluation quality. This paper chose Contribution Analysis (CA) as a methodological approach to study uses and influences of evaluations. We chose the qualitative approach combining different research techniques: documentary analysis, semi-structured interviews, direct observation. Two further steps were taken in this study: first, the intervention theory of change was described and second, translation theory was applied to analyze the movements of intermediary actors involved in the intervention. The use of CA with knowledge translation theory in this article, establishes a new methodological approach for utility and influence studies, supporting new developments referred to utility's standards in metaevaluations. It is presumed that this hybridization helps constructing participative evaluations models, since: a) emphasizes movements and mechanisms on intermediate actors connections and relationships trough translation operations, essential for understanding power relations on decision making; and b) responds to the necessity that evaluations have to inform action with legitimate and valid evidences. (Mayne, 2008). Thus, study has potential to build evaluations models that prioritize the utility and influence on decision making and on program improvement. Moreover, it may respond to expectations of evaluators who have interest on enhancing the meaningful use of their evaluations and in development of evaluation theory. Findings show the need to rethink utility's and influence's attributes, systemizing evaluations process that seek organizational knowledge, and enhance analytical alternatives on evaluation's field.

**References:**

JANUZZI, P. M (2001). Social indicators in Brazil: concepts, data source and applications. Campinas: Alinea.

Mayne J (2008). Contribution Analysis: An Approach to Exploring Cause and Effect. ILAC Brief No. 16: Rome: The Institutional Learning and Change Initiative. URL: [http://www.cgjarilac.org/files/publications/briefs/ILAC\\_Brief16\\_Contribution\\_Analysis.pdf](http://www.cgjarilac.org/files/publications/briefs/ILAC_Brief16_Contribution_Analysis.pdf).

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 016 Values and Human Rights Issues

## O 048 - Ethics in Evaluation; A Factor of Resilience? Whose Resilience?

*F. Etta*<sup>1</sup>

<sup>1</sup> Africa Gender & Development Evaluators Network -AGDEN, AGDEN, Nairobi, Kenya

What, if anything does the 'Me Too movement' tell us about the resilience of institutions, nay traditions on one hand and of women as other oppressed in the workplace on the other? What about the story which broke in the Times of London of Oxfam officers' sexual harassment and abuse of women and minors in Haiti in 2011? In the words of Sean O'Neil, who broke the story, there was 'alleged bullying, harassment, intimidation of Haitian and international staff and serious sexual misconduct – '; <https://www.thetimes.co.uk/edition/news/top-oxfam-staff-paid-haiti-quake-survivors-for-sex-mhm6mpmgw>.

What do ethics in evaluation have to do with this? Are the ethics in evaluation enough?

The objective of this paper is to bring these issues to the fore and ignite discussions and conversations among evaluators; on issues of; persisting moral challenge in development action which is said to be related to gender and power inequalities, cultural breakdown, changing norms and mores, wars, political instability and impunity; and the role and place of evaluation in support of gender equality and social equity and ultimately durable and sustainable societal resilience.

The paper argues that now is a good time, like the Me-Too movement, for evaluation to take head on and to address the pervasiveness of gender inequality and social equity. Using the extensive work done by the Africa Gender & Development Evaluators Network (AGDEN), a proposal is made for practical strategies for how moral, ethical and rights principles and standards of practice can be translated into competencies and actions possible with and within each evaluation whether at the national, institutional, programme or project level.

Using the AGDEN approach (developed over the last decade and a half), the paper presents propositions distilled from the work of the network for programme development actors as well as evaluators and one law for gender and rights responsive transformative evaluation. The paper holds that the notion that the ethics of research and/or of evaluation can by themselves deal with the problem is unsubstantiated because of the over reliance on methods and standards and the non-critical way evaluators apply these in Africa. The four propositions and one law are deployed to show how evaluation can be transformative of social and gender relations and how the results might be supportive of resilience among oppressed groups in society.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 016 Values and Human Rights Issues****O 049 - The Other Side of Gender Inequality: Men and Masculinities in Afghanistan**L.W. Echavez<sup>1</sup><sup>1</sup> University of Bohol, University Research Center, Tagbilaran City, Philippines

Masculinity refers to how people perceive the characteristics and qualities associated with being men. The overall purpose of the research is to achieve an in-depth understanding of different notions of being an Afghan man, and the constructs of masculinity in terms of roles of men and women, equality, power and control that contribute to gender inequality in three different study areas of Afghanistan including *Nangarhar*, *Takhar*, and *Bamyan*, and Kabul provinces. The study seeks to undertake both qualitative and quantitative approaches aiming to explore the meaning and understanding of being an Afghan man. It made use of the quantitative method utilizing the *Mann-Whitney U*-test in comparing the results of young and mature men and women, and *Kruskal-Wallis H* test if there is a significant difference among answers according to provinces. Qualitative methods through the in-depth interview, key-informant interview and focus group discussion in gathering data in the research sites. Findings showed that majority of male and female respondents among the four provinces are in agreement on the various masculine stereotypes on roles of men and women; equality, control and power; and gender-based violence. Result revealed a plurality of views concerning wife battery and no significant variation between the responses of young and mature respondents. Acceptance of the practice of wife beating is highest in Nangarhar and lowest in Bamiyan. A little less than half subscribed to the traditional views of Violence Against Women and Children, the highest percentage of those came from Nangarhar. Such findings are apparently validated in the qualitative part. Majority of the respondents agreed that gender-based violence although not right is justifiable when women resist the decisions made by the men. Most of the male religious key informant believed that basing on Islamic rules, it is permissible to beat a wife in case she is a "*Nashiza*." As validated in the fgd, such term means rebellious/disobedient woman. In this circumstance culture becomes a facilitator, at the same time, a barrier to change. The findings will contribute to deeper comprehension among various types of Afghan masculinities and how these patterns involving gender schemas impact gender inequality in Afghan communities.

**Keywords:** Gender and Development, masculinity, Quantitative- Mann-Whitney U-test, Kruskal-Wallis Test, Qualitative methods- FGD, KII, IDI, South Asia, Afghanistan

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

## S 016 Values and Human Rights Issues

### O 050 - Women Leaders Project, Madrid City Council. An Evaluation Based on the Theory of Change With Gender & Human Rights Approach

*S. Franco Alonso<sup>1</sup>, A. Ballesteros Pena<sup>1</sup>, M. Donayre Pinedo<sup>1</sup>, P. Serrano Garijo<sup>2</sup>*

<sup>1</sup> EMDH - Education Multiculturalism Human Rights, Evaluation, Madrid, Spain

<sup>2</sup> Madrid City Council, Dirección General de Personas Mayores y Servicios Sociales del Ayuntamiento de Madrid, Madrid, Spain

This article presents the key elements of the evaluation of the project "Promotion of the participation of senior women: Women leaders of the Villaverde District" in Madrid, promoted by the Madrid City Council between 2014 and 2017. The objective of this project is to increase the participation of elderly women in the Local Centres for Senior Citizens and beyond, from a gender, empowerment, active ageing and human rights approach.

The methodology of the evaluation carried out has been based on the theory of change approach, which has made possible to identify how and what results were intended to be obtained. For this purpose a theoretical model has been developed, on which the evaluation process and the analysis have been articulated. This approach has incorporated the gender and human rights perspective as well.

The fieldwork has involved the application of the following evaluative tools: an in-depth documentary review, more than thirty interviews with key stakeholders, four participatory workshops with elderly women leaders and other Local Centres for Senior Citizens members, and non-participatory observations in nine different relevant project-related activities.

On the one hand, the findings resulting from this evaluation process have made possible to identify the potential that empowerment processes have in the population of senior women from diverse axes: personal, social and political. On the other hand, the evaluation has highlighted the need for the gender perspective to be incorporated in an articulated manner by all the actors involved in the interventions, with the use of methodologies that respond to the needs and experiences of the target populations.

A series of key issues emerged throughout the evaluation process, in terms of initiatives to mainstream gender and human rights principles in the field of working with senior women. These include: the need for interdepartmental coordination, the essential presence of gender-sensitive and human rights-sensitive actors, the commitment and support of all the institutions involved and at all levels (political and technical), the relevance of giving people a leading role in promoting their autonomy, and the importance of incorporating the principles of active ageing and also to promote the right to participation.

Based on the conclusions of the evaluation, a set of good practices and key lessons learned have been identified for its future replicability in the rest of the districts of Madrid. Likewise, a series of recommendations have been proposed, which can be used for the design and implementation of future projects of similar nature and objectives, that is, in projects focused on the empowerment of senior women, active aging and human rights.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 017 Policies, Governance - Influence of Evaluation**

**O 051 - Institutionalisation of Policy Evaluation as Enabler for Sound Public Governance: Towards an OECD Perspective**

*E. Beuselinck<sup>1</sup>, I. Stola<sup>1</sup>, S. Schmitz<sup>1</sup>, D. Milan<sup>1</sup>*

<sup>1</sup> OECD, Public Governance Directorate, Paris, France

This paper presents the OECD's ongoing efforts to develop a comprehensive view on the institutionalisation and use of policy evaluation across OECD countries. Both institutionalisation and use of policy evaluation can be considered as important indicators of the maturity of a country's policy evaluation system, enabling evidence-informed policy making and – therefore – sound public governance.

While several OECD countries are progressively reinforcing or reforming their policy evaluation systems, few normative claims exist on how to embed these in countries' existing governance architecture (Jacob et al., 2015). Systematic, internationally comparative stock-taking exercises are limited[1], but it is clear that the institutional set-up of policy evaluation varies significantly across countries. A Prime Minister's Office, an independent agency, the Ministry of Finance or a horizontal, theme-oriented ministry (e.g. 'social impact') can all provide the main institutional anchorage for policy evaluation. Moreover, as the institutional set-up differs among countries, so does the underlying rationale for institutionalisation. A country's political system, champions inside or outside the Executive, requirements issued by international or supranational organisations, fiscal pressure, or a vibrant evaluation society can push for policy evaluation and its institutionalisation (Gaarder and Briceño, 2010).

Taking the aforementioned elements into consideration, the OECD Public Governance Directorate has decided to develop and launch in 2018 a survey on policy evaluation across OECD Members and selected non-member countries. This stock-taking exercise will address amongst others countries' legal and policy framework for policy evaluation; key actors; availability and nature of policy evaluation guidelines; perceived objectives and challenges; quality assurance efforts and mechanisms to promote the use of policy evaluation.

The paper will present the OECD's approach to foster the understanding of the nature, degree and underlying rationale of institutionalisation and use of policy evaluation. For that purpose, the paper will address the following issues: (1) main features of the survey (e.g. type of respondents, key components of the survey, lessons learned from the piloting phase); (2) next steps for data collection and analysis; (3) ongoing and envisaged initiatives to complement the comparative survey data with more in-depth country work; (4) priorities for future work. Moreover, the paper will also reflect upon the potential contribution of this work to the further strengthening of a policy evaluation community across OECD countries and the relevance of developing OECD standards to advise governments in the development of robust policy evaluation systems, adapted to their specific needs and political, historical and institutional reality.

[1] Important recent and ongoing initiatives include the International Atlas of Evaluation by Furubo, Rist and Sandahl (2002), updated by Jacob, Speer and Furubo in 2015; and The Evaluation Globe – Compendium on the World-Wide Institutionalization of Evaluation launched by the Centre for Evaluation, Saarland University.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 017 Policies, Governance - Influence of Evaluation**

**O 052 - The Influence of Evaluations in Polycentric Settings:  
a Literature Review**

*M. Crijns<sup>1</sup>*

<sup>1</sup> Open University of the Netherlands, assistant professor, Heerlen, Netherlands

The use and influence of evaluations is a core topic for evaluation theory and practice, but few empirical studies address the mechanisms that generate the influence of an evaluation. The contextual characteristics of the decision or policy setting in which evaluations are conducted have evolved over time. The setting of various policies has been characterized more and more by polycentric features, such as the involvement of multiple governing actors at different levels rather than a monocentric hierarchy, and considerable independence of each actor to make norms and rules within a specific domain (Ostrom, 2010). Climate governance is a prominent example of polycentric governance, as Elinor Ostrom argued.

An important question arises: since the policy context nowadays of a lot of programs and policy under evaluation is evolved to a polycentric setting, is this setting recognized and taken into account in the evaluation influence literature?

The purpose of the study is to find out if polycentricity may affect the conditions for the influence of evaluations. To achieve our purpose we undertake a systematic review of the empirical literature on evaluation influence. We aim at answering the following research question: *Is the polycentric setting, in which much of current policy takes place and is evaluated, problematized in the evaluation influence literature as context factor that could play a role in the influence of evaluation on policy or policy implementation by organizations and individuals? If the polycentric setting is problematized, how does the literature say it needs to be addressed? If not, which role has the polycentric setting in fact played in evaluation influence yet, although not problematized as such?*

Preliminary findings point to a gap in the evaluation influence literature: the polycentric setting is seldom problematized, and therefore has not been studied as a mechanism that plays a role in the influence of an evaluation. However, in studies showing indications of polycentricity we try to analyze which role a polycentric setting in fact plays in an evaluation and its influence. The role of management (gatekeepers) and evaluator or evaluation team (knowledge broker) and their mutual relations, also with stakeholders, seem to be important for that matter.

The results of the literature analysis will suggest avenues for future research on the influence of climate policy evaluation.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 017 Policies, Governance - Influence of Evaluation**

**O 053 - Can Evaluations Really Contribute to Evidence-Based Policy Making at Government Level? – The Case of the French Government Modernisation Evaluations**

*T. Delahais<sup>1</sup>, L.F. Clément<sup>2</sup>*

<sup>1</sup> *Quadrant Conseil, PARIS, France*

<sup>2</sup> *Ministère de l'Action et des Comptes publics, Direction interministérielle de la transformation publique, Paris, France*

It is a common-sense observation that policy making does not always rely on adequate information. Of course, political choices do not need to follow what evidence apparently dictates. Yet the continuing debate also proves that even the most consequential decisions may not be 'evidence aware' (Nutley & Webb, 2000).

Evaluators are often concerned about the use of evaluation in policy making, as the amount of empirical research on this topic tends to suggest (Henry, 2004). This is because evaluation is a source of evidence that is in many cases directly aimed at informing decision-making (Patton, 2008). Given the many factors influencing decision making, how can we accurately trace evaluations' contribution to this process?

To answer this question, we build upon an evaluation commissioned in 2016 by the French Secretary General for Modernisation of Public Action, SGMAP, which included *inter alia* a metaevaluation aimed at assessing the quality of 65 evaluations launched between 2013 and 2016 by the French Government, and an 'evaluation of evaluation' using a contribution analysis (CA) approach focused on a subset of 8 evaluations. The metaevaluation part was based on a comprehensive documentary analysis, including previously undisclosed administrative decision-making documents related to each of the 65 evaluations. The evaluation of evaluation relied on a case-study approach, including additional documentary analysis and 59 semi-structured interviews with evaluation stakeholders, within and outside the French State administration. One of the authors was involved in delivering the evaluation services while the other steered this evaluation as an SGMAP officer.

The examples of metaevaluations in France with a focus on actual use are few and far between (with the notable exceptions of (Toulemonde, Genard, Jacob, & Varone, 2006) and (Epstein, 2009)). This study thus provides a unique empirical material to analyse both the use of evaluation in the French context and contribute to the methodological debate about the actual influence of evaluation on policy making.

The paper first lays out the (incomplete) institutionalisation process of evaluation in France that led to the Modernisation de l'action publique evaluation programme, and analyses how it supposedly should have affected the role of evaluations in decision-making. Then, using the metaevaluation it identifies the attributes of the evaluation process and reports that should, as per our theoretical framework, lead to stronger use in decision-making, and assess the 65 MAP evaluations accordingly. Finally, it assesses the actual contribution of these evaluations to decision-making, zooming on a smaller set of 8 case studies and using a set of empirical tests inspired by Process Tracing used in combination with Contribution analysis to assess a contribution.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

## S 018 Evaluating Sustainable Development and 2030 Goals 1

### O 054 - Why Does Europe Leave SDG Evaluation Behind?

*W. Meyer<sup>1</sup>, B. Simon<sup>2</sup>, P.D.R. Stockmann<sup>3</sup>*

<sup>1</sup> Saarland University, CEval, Saarbrücken, Germany

<sup>2</sup> Planete Publique, Sustainable Development, Paris, France

<sup>3</sup> Center for Evaluation, Saarland University, Saarbruecken, Germany

Agenda 2030 (so called Sustainable Development Goals) which were adopted by the UN General Assembly in 2015 provides a relevant, systemic and inclusive framework to build up resilient societies thanks to an effort to combine major stakes into a new development model.

To succeed in its achievement, reporting and accountability are key as expressed in article 47 of the UN declaration: “Our Governments have the primary responsibility for **follow-up and review**, at the national, regional and global levels, in relation to the **progress** made in implementing the Goals and targets over the coming 15 years. To support **accountability** to our citizens, we will provide for **systematic follow-up and review** at the various levels, as set out in this Agenda and the Addis Ababa Action Agenda. The **high-level political forum** under the auspices of the General Assembly and the Economic and Social Council will have the central role in overseeing follow-up and review at the global level.”

In this perspective Voluntary National Reviews (VNR) are annually presented by UN Member States at the High Level Political Forum, reporting on their level of achievement of SDGs and thus displaying their evaluation framework. Especially the first VNR of a country is supposed to give an overview on its reporting system, including Monitoring and Evaluation of SDGs as a significant element. Being in the third round of VNRs, it is time for some stocktaking on a total of almost 80 countries in the World. Almost all European countries – exceptions are Austria and UK – are included in this list. The most significant finding of research on the inclusion of Monitoring and Evaluation of SDGs into national SDG reporting is the poor performance of European countries compared to the rest of the World.

The purpose of this paper is an attempt to explain why European countries seem to leave evaluation behind for the implementation of SDGs in a three part analysis:

- an exhaustive screening of the content of the 2016, 2017 and 2018 VNRs from an evaluation perspective thanks to an in-depth document analysis and key-words;
- a comparison with low income countries to appreciate the relative effort of EU countries;
- last a deeper view on 6 European countries to find key elements of explanation why countries with a relatively high evaluation culture are lagging behind in the framework of SDGs on the basis of national strategic documents and interviews with key national stakeholders.

We will then stress few recommendations for a stronger use of evaluation for the implementation of SDGs at the national level.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 018 Evaluating Sustainable Development and 2030 Goals 1**

**O 055 - SDGs National Reports, the Gap Between the Reality and What Governments Show. A Review of Latin-American Reports About SDG 5**

*K. Paez*<sup>1</sup>

<sup>1</sup> *Chilean Evaluation Network, Evaluation of SDGs, Santiago, Chile*

This paper is a systematization that analyzes problems which appear during monitoring the SDG 5, incorporating in its reflection not only the compliance status from its indicators, but also investigates existing conditions for its monitoring and makes a reflection on the possibility of integrating other indicators relevant for Latin American context to support their follow-up.

The methodology used is mixed. First, it is focused on the quantitative methodology, based on the collection of secondary data, where a parallel is made between the information provided in the official reports of Latin-American reports and the available public information used in the civil society's report.

The main results of this exercise have been the discrepancy between the information provided by the States in the voluntary report and the existing data, since in this first report there are more data available than those that the State delivered. At the same time, it is recognized that new process indicators must be incorporated to improve monitoring compliance with the associated goals, and finally it has been established that standards are a key tool in monitoring this SDG from civil society, where the State is still in debt in their application.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 018 Evaluating Sustainable Development and 2030 Goals 1**

**O 056 - Sustainability Reporting, SDGs, UN-Global Compact and GRI: How is Private Sector Communicating its Impact?**

*M. Branco*<sup>1</sup>

<sup>1</sup> [www.mariana-branco.com](http://www.mariana-branco.com), EES Board- EES Emerging Evaluators- EES Social Media, Porto, Portugal

**Justification:** Human development demands for partnerships and co-creation. To build resilient societies all actors must be engaged: private, public and social sectors. Moreover, different publications provide clear evidence that prioritizing sustainability can increase market share and sales. Recognizing this the private sector engagement and investment in sustainability is increasing. Many companies are searching for answers on how they can maximize the social impact of their investments and how to communicate this impact.

**Rational:** This paper will focus on the second question – sustainability communication. More concretely, on how companies struggle to communicate the success of their social actions in an ecosystem where new referential and standardization attempts pop up at every corner. It will explore this topic, sharing 2 years of experience analyzing the communication towards sustainability of 16 portfolio companies, such as Unilever, Danone, Nestle, Lindth, HSBC, Essilor, Geberit, Fresenius and Publicis.

**Objectives:** This presentation will debate the following topics:

- Importance of Global Compact
- Importance of combining Global Compact with GRI
- Good materiality analysis
- Increase use of theory of change
- Increasing use of good story-telling
- Lack of reference to conflict prevention
- Lack of outcomes measurement
- Scarce impact evaluation
- GRI not serving as a basis to demonstrate impact
- SDGs are important as a theoretical background by they are not useful to make decisions
- Lack of reference to UN guiding principles on business and human rights

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 019 Science Technology Evaluation

## O 057 - Using a System Approach to Rationalise Policy Evaluation in the Field of Science Communication and STEM (Science, Technology, Engineering, Mathematics)

*M. Van Hoed*<sup>1</sup>

<sup>1</sup> IDEA Consult, Research and Innovation, Brussels, Belgium

Since 1993, science communication has gained importance in the Flemish Science and Innovation policy field. From the start, the aim was to improve and support the communication of scientific activities and outcomes to the broader public and to thus increase the societal support for science and technology. The Flemish government supports a variety of small to medium actors and bottom-up initiatives in the field, some in a structural way, others on a project basis.

In the past decade the fields of Science, Technology, Engineering and Mathematics, combined in the concept of 'STEM', have become an increasingly important focus in science communication. The introduction of STEM was triggered by an increasing need for graduates in exact scientific and technical fields and required a stronger cooperation between the policy domains of 1) Education and Training, 2) Work and Social Economy and 3) Science and Innovation.

In the context of this complexity of initiatives across policy domains and the evolving policy context, the government requested a combined set of evaluations of several of the science communication actors in Flanders and one systemic evaluation. The aim of the evaluation was:

- to evaluate the role, functioning and impact of the individual actors in the field, as well as their contribution to the broader policy objectives;
- to evaluate the system of science communicators in Flanders, their interaction, cooperation and value added in view of their contribution to the broader policy objectives;
- to formulate recommendations for the rationalisation of the system of science communication in Flanders.

The evaluation took both an ex-post and ex-ante perspective. New is the systemic view taken in the evaluation, within the complex context of multiple actors, initiatives and policy domains involved. It was important to take into account the policy objectives of the different domains and at the same time address other policy-relevant issues such as the effectiveness of local versus centralised initiatives and the complementarity of different actors in view of reaching the broad target audience, and in particular also gender equality in STEM.

Also the policy learning aspect was emphasised. At the level of individual actors and initiatives, the broader public was consulted in a survey to assess their reputation and use, but also their impact on opinion building and intake in education in STEM-fields. At the level of the science communication system, the government asked explicitly for a strong SWOT-analysis, development of monitoring/KPI and recommendations for rationalisation of the system. In this view, stakeholders were interviewed and a validation workshop with stakeholders from the entire system was organised.

We will present the set-up and outcomes of this evaluation, paying specific attention to the lessons learnt from this complex setting and the systemic evaluation approach. We will further explain how this approach impacted on policy learning and the further development of the science communication system in Flanders towards a mature system, touching upon many policy fields and horizontal issues, and contributing to a stronger and more resilient society.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 019 Science Technology Evaluation

## O 058 - Assessing Impact of Taiwan's Publicly-Funded STI Program on Participant Firms: the Empirical Evidence and Policy Implications

C.C. Chang<sup>1</sup>, T.Y. Liu<sup>2</sup>

<sup>1</sup> Science and Technology Policy Research and Information Center- National Applied Research Laboratories, Policy Research Division, Taipei, Taiwan

<sup>2</sup> Science and Technology Policy Research and Information Center, Policy Research Division, Taipei, Taiwan

Utilization of evaluation in further policy making has long taken center stage regarding formative evaluation development (Scheirer, 2012; Glasgow et al., 2014; Howell & Yemane, 2006; Jordan, 2015). However, the feedback loop of program-level formative evaluation might be systematically obstructed by the institutional context regarding Taiwan's STI policy cycle. The pressing institutional issues such as insufficient utilization of reported quantitative performance data due to funding agencies' inadequate evaluation capacities and managerial capabilities need to be addressed in addition to the summative impact assessment.

Therefore, this evaluation research is designed to positively measure the output additionality, behavioral additionality and project additionality by means of econometric methods, and then explore the institutional causes embedded in funding agencies' program planning/implementation/evaluation practices.

Before designing the impact evaluation, we conducted the evaluability assessment to employed by many STI funding agencies (e.g., NIH, DOE, etc.) to assure the readiness of summative impact assessment and identifying the non-alignment of the monitoring indicators with the managerial needs with newly constructed logic model to guarantee the utilization of evaluation results.

In terms of quantitative impact assessment, we first simultaneously assess the output additionality of Taiwan's STI Project for Intelligent Electronics) combining the DID regression and Matching Method. When assessing the output additionality of NPIE, we combined the DID with Matching Method (exact matching) to exclude the specific firm characteristic variables (related to selection bias) and unobserved variables (Caliendo & Kopeinig, 2008; DIDE-CUMSPH, 2013). We employ the DID regression method to examine the causal relationship (Dimick & Ryan, 2014; DIDE-CUMSPH, 2013) and use semi-parametric DID to determine the average effect of program intervention on the treated firms (Houngbedji, 2015).

Secondly, combining the intermediate impact composition mentioned above on participant firms, this study chose to measure those behavioural additionality including "Creating inter-organizational linkages", "Improving quality or sale of products and processes", "Enhancing R&D or innovation capacities", "Enhancing capability of commercializing technologies", "Enhancing technological expertise of employees", "Increasing employment of high-quality personnel", "Reducing R&D/production/operation cost", "Broadening the scope of R&D", and "Facilitating technology adoption" through client-based survey.

Thirdly, since the intervention regarding subsidizing private firms must be justified by the underinvestment of R&D investment due to appropriability problem (Gomper & Lerner, 1999), we separately assess the impact of NPIE on subsidized firms in terms of input additionality (increased R&D intensity), project additionality, scope additionality, and acceleration additionality (Lohmann, 2014).

Fourthly, based on the negative assessment results regarding output additionality and project additionality, we compared the quasi-experimental impact assessment results from this research and from other similar impact assessment (Hsu et al., 2009) with the previously reported quantitative outcome indicators, thereby finding the disagreement between them.

Finally, we conduct the qualitative interview to identify the misleading effects of quantitative indicators which wrongly incentivize the program staff to form the conservative R&D program which firms would have been willing to do absent the program intervention. Through this impact assessment, we try to provide quantitative evidence to navigate the refinement of Taiwan's institutional structure regarding the STI program planning, monitoring and evaluation.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 020 Evaluation of Training Programs**

**O 059 - When the Measurement Explains the Effect: The Effect of Baseline Survey Participation on Intercultural Training Effectiveness Outcomes**

*M. van Egmond<sup>1</sup>, B. Froncek<sup>1</sup>, V. Piper<sup>1</sup>, A. Rohmann<sup>1</sup>*

*<sup>1</sup> University of Hagen, Community Psychology, Hagen, Germany*

The influence that demand characteristics, measurement effects and response biases have on evaluation outcomes is well-known in the social scientific and evaluation literature (e.g., Podsakoff, MacKenzie, Lee & Podsakoff, 2003). Yet, the question whether the mere act of participating in an evaluation survey might influence training outcomes has not been examined extensively yet. In the evaluation of an intercultural training, we found that participants' intercultural sensitivity (Fritz, Möllenberg, & Chen, 2001) significantly increased not as a function of participation in the training alone, but as a function of participation in repeated measurements of intercultural sensitivity as well. The study was conducted at a police department in the lower-Saxony region of Germany. A waiting-control group design was applied and the full sample consisted of 175 participants. Analyses of variance and hierarchical linear regression analyses revealed that the intention to behave in interculturally sensitive ways, perceived behavioral control as well as interaction attentiveness in intercultural situations increased only among training participants in the waiting control group, who had participated in two Baseline surveys before training exposure. Moreover, in both the waiting control and the training-group, after three measurement points, an increase on the dimensions of interaction engagement, respect for cultural differences, and diversity beliefs occurred – irrespective of participation in the survey or in the training. The implications of these findings for evaluation designs that rely on self-reflection and self-awareness of evaluation participants will be discussed.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 020 Evaluation of Training Programs

## O 060 - Evaluation of Non-Formal Education for Mental Health Professionals: The Case of the Appreciative Systemic Inquiry Experiential Educational Program of OKANA

A. Kitsiou<sup>1</sup>, P. Karakoula<sup>1</sup>, I. Gotsis<sup>2</sup>, A. Kiapoka<sup>2</sup>

<sup>1</sup> Greek Organization Against Drugs OKANA, Research and Evaluation, Athens, Greece

<sup>2</sup> Greek Organization Against Drugs OKANA, Education and Human Resources, Athens, Greece

Education and vocational training quality improvement is fundamental for educational policy both in national and EU level, while educational policy making is now based on research data and evidence gathered from quality assurance procedures. Although evaluation has been mostly related with formal forms of education, the importance of evaluation for non-formal education has been also emphasized in the last two decades.

Regarding adult education, in particular, and according to OECD guidelines, evaluation is considered increasingly necessary for professionals' improvement, in order to be adapted in their organizations' needs, accordingly to the modern economic, social and political developments. However, in Greece, there is still lack of evidence for the quality assurance methods -and their results- utilized in the non-formal types of vocational education.

Up to this point, the present study presents the quality assurance and evaluation methods utilized for the experiential educational program in Appreciative Systemic Inquiry (ASI), a non-formal education program, designed and implemented by the Education Department of OKANA, and addressed to mental health professionals. Given that the evaluation for each program constitutes a specific process with specialized characteristics, we employed a pilot method based on Kirkpatrick's four level assessment, and utilized quantitative and qualitative instruments (triangulation). The instruments include **a. a structured questionnaire**, completed before the program's beginning, assessing the participants' expectations from it, related to their professional needs, **b. reflective diaries**, written after each educational section, exploring their personal experience and gradual movements during the program, **c. focus groups**, conducted after the completion of the program, assessing its typical characteristics and the participants' personal experience, **d. a structured questionnaire** for the overall assessment of the program's typical characteristics, and **e. a follow-up questionnaire**, completed 6 months after the completion of the program, evaluating the program and the utilization of the acquired know-how in the participants' work field.

Research findings are expected to contribute to updated data regarding the evaluation of non-formal vocational education addressed to mental health professionals, and form the ground for a prolific dialogue on the necessity for the evaluation of non-formal education programs, which education policy makers should take under consideration, since these programs constitute an integral part of the modern society of knowledge and lifelong learning.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 020 Evaluation of Training Programs**

**O 061 - Is Training Evaluation a Threatened Discipline?  
The Dilemma of Theory Vs. Practical Reality**

*C. Duffy<sup>1</sup>, S. Goodman<sup>1</sup>*

<sup>1</sup> *University of Cape Town, School of Management Studies, Cape Town, South Africa*

Theorists and evaluation experts argue the importance of training evaluation in the training cycle, yet there is little evidence of evaluation application across the globe. Research from both international and local studies has demonstrated that organisations are not adequately evaluating their training efforts, with the practice being limited to the collection of reaction level data. In this paper we argue that without a radical shift in how training evaluation is conceptualised as well as the development of new approaches / methods, the discipline will cease to exist in the country.

The paper reports on the first study investigating training evaluation practices across corporate South Africa and explores the opinions held by top human resource (HR) executives about this training and business practice. Results from the research mirrored those of previous global training evaluation trends. The South African corporates sampled are not conducting objective training evaluations but rather use idiosyncratic mechanisms and methods to monitor and report on their training. There is also no genuine willingness to determine the effectiveness of their training and development investments. Training evaluation is not viewed as a worthwhile business practice. Opinions expressed suggest that training evaluation is a waste of time and there is little prospect of this perception changing without a significant departure from how training evaluation has been theoretically positioned and practically operationalised.

The research emphasizes a disconnect between theory and practice. Traditional training evaluation methods and models are outdated. They appear futile in the highly demanding corporate environment, and impotent given the way organisational learning is taking place. We recommend a substantial shift in how training evaluation is conducted. Context should inform practice. For organisations to perform training evaluations, evaluators must rethink training evaluation design, approaches, and methods. Evaluators must work with what people are doing in their current HR environments and find solutions to enable them to extract greater understanding of the training they provide. Only then will corporates begin to engage meaningfully in training evaluation and not waste opportunities for organisational learning and change.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 021 Learning From Evaluation

## O 062 - Resilience Through Adaptation: The Role of Evaluation in Driving Reflection and Change in Adaptive, Democracy Assistance Programmes

*G. Ramshaw*<sup>1</sup>

<sup>1</sup> *Westminster Foundation for Democracy, Director of Research and Evaluation, London, United Kingdom*

Clearly defined objectives and a well crafted theory of change are critical to conducting results-oriented programming. But evaluators and implementers alike also understand that to achieve its goals in complex environments, a program must respond effectively to emergent challenges and opportunities. Rather than force programs to comply with an increasingly outdated results framework, properly designed and implemented monitoring and evaluation tools can encourage so-called "Adaptive Management." These tools deliberately incorporate processes for programmatic learning and adaptation in response to a changing context. In doing so, they can encourage resilience both in terms of the programme itself and the broader societal outcomes it hopes to sustain.

Adaptive management practices are intended to encourage organizational learning throughout the program lifecycle, empowering organizations to deviate from a theory of change and results framework if evidence suggests that these deviations will improve the program's ability to meet stated goals and objectives. While it is sometimes contrasted with a more traditional "results-based management" approach, adaptive management is most effective when it follows a deliberate, informed process for making decisions throughout the program lifecycle. Outlining – and following – this process is a key hurdle to implementing an effective adaptive management approach. Because adaptive management envisions a continuous feedback loop between design, implementation, monitoring and evaluation processes, it requires coordination and communication between funders, program managers, technical/country experts, field teams, evaluators, even beneficiaries. Organizations that work in the democracy and governance sector have long grappled with the challenges of conducting programs in complex, fast-moving environments that require continual adaptation.

Integrating these new adaptive management practices requires more than just the creation of new tools. It entails a redefinition of the role that evaluation often plays in organisations, particularly in democracy assistance. Evaluation must step beyond its role as accountability agent or post facto results collector and become an equal partner in the programme design and management process. Resilience through adaptation requires a renegotiation of the power dynamics within organisations to enable evaluation to contribute its evidence at key decision points in the programme management cycle. It is often said, "It is not the strongest species that survive, nor the most intelligent, but the ones most responsive to change." Evidence from our evaluations and organisational learning that we will present in this paper suggests that resilience is born from evolution, not steadfastness. Evaluation has a key role to play in ensuring that organisations in democracy assistance and other fields are able to achieve resilience by being adequately responsive to change.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

## S 021 Learning From Evaluation

### O 063 - Knowledge-Focused Evaluation in Changing Contexts – What Evaluation Synthesis can Bring to the Table

*J. Pennarz*<sup>1</sup>

<sup>1</sup> Director- Independent Office of Evaluation of IFAD, Lead Evaluation Officer, Rome, Italy

The context for international organisations is changing in rapid, complex and unpredictable ways. Development organisations are under pressure to adapt their business models, to address the challenges resulting from reduced donor funding and to meet the ambitious developmental targets under the Agenda 2030. Evaluation has an important role to play, facilitating learning loops and supporting adaptive change. We have seen in our organisations an increasing appetite for knowledge products that systematise evidence and lessons from evaluations for organisational learning.

The Independent Office of Evaluation (IOE) of IFAD prepares evaluation syntheses that are intended to address emerging demands for evaluative knowledge. Evaluation syntheses are desk-based studies that capture the accumulated knowledge from existing evaluative evidence on a specific topic and identify relevant lessons to enhance development effectiveness. They cover a broad range of issues that are of interest for upcoming strategic or policy planning in IFAD, but which are equally relevant to the development field, in general and are therefore widely shared

The presentation will discuss the opportunities and challenges related to evaluation syntheses. Synthesis reports draw from a wide range of evaluation sources and are thus able to generalise and provide with broader insights. As a knowledge product they go beyond presenting evidence from evaluations; they also provide state-of-the art concepts as well as stakeholder knowledge in order to contribute to the on-going strategic thinking. To be useful, syntheses are expected to link the evaluative evidence with the challenges of the future. In other words, not only do they ask: what has worked in the past (and why)? but also: what will work in the future (and under what conditions)?

The presentation will strategies to address these questions, depending on the nature of the topic. Strategies include (1) applying a "fit for the future" lens to the analysis of documented practices and results; (2) adaptive search for evidence to squarely address the limitations in the available sources; (3) generalisation through hypothesis-led inquiry and use of typologies; and (4) contextualisation through case studies and qualitative analysis.

These strategies will be discussed using the examples of four synthesis reports: gender-transformative practices and results, inclusive rural finance, partnerships, and scaling up. The synthesis on gender (2017) developed a forward-looking concept to analyse the existing evaluative evidence: Transformative change is at the heart of the Agenda 2030 and the conceptual framework used to systematise gender-transformative practices and results has been widely perceived and discussed.

The presentation intends to contribute to the on-going debate on (a) how to make existing evaluative evidence accessible and usable for decision makers given the dynamic context in which it is collected and in which it is to be used; and (b) if findings from evaluations were to focus on external validity or contextualisation in order to be useful. The experience presented here suggests that in order to be useful, the synthesis has to be knowledge-oriented and rooted in the context.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 021 Learning From Evaluation**

**O 064 - Evaluation as Therapy: An Evaluation of Corporate –  
NGO Partnership between C&A, C&A Foundation and Save  
the Children**

*L. Risby<sup>1</sup>, I. Martins de Souza<sup>2</sup>*

<sup>1</sup> C&A Foundation, Effective Philanthropy, Zug, Switzerland

<sup>2</sup> Ar & Terra Consultancy, Evaluation, London, United Kingdom

The complex issues and questions of our times cannot be solved by evaluators writing a report, which may or may not be used by stakeholders and then finishing the work.

Learning is only meaningful if stakeholders act on the information that an evaluation can bring. For that to happen, they have to be part of the evaluation itself, rather than have the evaluation done to them.

In this paper, the authors – a former World Bank evaluator and a former ExxonMobil engineer – explore how the Whole Person Approach borrowed from psychotherapy was used to transform an evaluation from an extractive and often threatening exercise into a type of individual and group ‘therapy’ that provided enhanced opportunities for reflection, self-criticism, co-analyses and learning. Furthermore, it enables trust and confidence to be built between the external evaluators and stakeholders, keeping the right relational dynamics to improve use and utility of evaluation findings, and blurs the division between the role of the evaluator and the role of the subject being evaluated, whilst still maintaining the evaluation’s independence.

Traditionally, consultants do all the thinking, designing and drawing conclusions; whilst the stakeholders know nothing of what the consultant is up to in his/her mind until final results are presented. In contrast, the Whole Person Evaluation adopts peer-based processes to facilitate the co-design and co-sensing of results with global and local teams. By emphasising non-linear and creative thinking, valuing awareness of feelings, the approach encourages people to imagine and reflect upon issues before jumping into conclusions. For example, the use of metaphors, such as comparison of Superman’s superpowers and the kryptonite to discuss strength and weaknesses, creates a relaxed and informal environment, which in turn leads to more meaningful and rich responses in the field visits, workshops and semi-structured interviews.

The Whole Person Approach was applied in practice during the evaluation of a complex corporate global humanitarian partnership between C&A (Fashion retailer), C&A Foundation (its corporate foundation) and Save the Children. At the end of the evaluation process, stakeholders took full ownership of the results (positive and negative) and the report was used as road map for action in the next phase of the partnership. The external evaluation team were invited to continue their engagement with the partnership to provide regular feedback and learning in a process closer to developmental evaluation. The report was not the end of an evaluation, but the beginning of a new evaluation – without the Whole Person Approach, the authors argue, this would not have been possible.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

## S 022 Turbulent Contexts and Evaluation

### O 066 - How to Strengthen Our Coordination for Resilience of the Poorest: A Theory of Change & Learning from the Ethiopian Experience

*C. Rodriguez-Ariza<sup>1</sup>, B. Taye<sup>2</sup>, A. Nigatu<sup>3</sup>, Y. Mossisa<sup>3</sup>*

<sup>1</sup> *FAO, FAO Ethiopia Resilience Team, Addis Ababa, Ethiopia*

<sup>2</sup> *Program manager, European Union DEVCO, Addis Abeba, Ethiopia*

<sup>3</sup> *FAO, FAO Ethiopia, Addis Abeba, Ethiopia*

**Rational:** As a response to the 2011 and 2014 droughts in the Horn of Africa, the European Union (EU) has developed the initiative "Supporting Horn of Africa Resilience – SHARE" to advance food security, sustainable agriculture and resilience in the Horn of Africa. As part EC SHARE program, the FAO's three years Project "Strengthening Institutionalized Subnational Coordination Structures and Harmonization Mechanisms", called EC SHARED project and implemented in Ethiopia by FAO and EU (1) at the federal level and (2) through the regional government partners at regional level, and contributes to strengthen the Ethiopian planning and MEAL in resilience coordination mechanisms.

**Objective:** This communication aims to describe the project design and implementation process and results. The key actions of this project for strengthening coordination mechanisms are: (1) Coordination Capacity Development for planning, monitoring and evaluation (2) Knowledge Management in Coordination Spaces and (3) MEAL supporting decision-making and (4) Gender Mainstreaming in Coordination Spaces.

**Narrative:** Given the Ethiopian context and institutional framework, the project has adopted a flexible planning and developmental evaluation approach.

Given the multiplicity of actors & sectors involved in Ethiopian resilience interventions, the coordination function is a key element for improving resilience. Key international development processes are intended to strengthen coordination (Paris Declaration kind, Collective Impact, Agenda 2030, New Way of Working...)

There are many different definitions of resilience, also there are many different definitions of coordination used in research and practice. This implies that there are still many opportunities and challenges for the combination of both (resilience coordination), some of them are: (a) understanding/agreeing on the concept of resilience coordination, (b) unpacking the theory of change of the resilience coordination processes and their drivers of change

In practice some of the key resilience coordination challenges are: (1) The fragmentation, limitations in linkages and synchronization of actors and the tradition to work in silos of the stakeholders, (2) The lack of strong coordination backbone organizations and coordination champions, (3) The gaps in credibility and trust between stakeholders, (4) The weak "real" incentives for collaboration and the strong "real" incentives for competing for funds or recognition.

The project has developed a framework of analysis based in a set of drivers for resilience coordination:

(1) The existing barriers for coordination are underestimated by the stakeholders and, in practice, existing incentives for coordination are insufficient for (a) overcoming these barriers (b) supporting coordination championship,

(2) FAO/EU coordination and collective impact approach comprise three steps: (a) Strong leadership, common vision and mutual accountability; (b) Common agenda and joint capacities/actions; (c) Strong system of incentives (monitoring & knowledge management / measurement systems, transparency, proper & reliable communication, trust, win-win activities & funds and gender mainstreaming),

(3.) Promoting a coordination mutual accountability framework is the key instrument for aid effectiveness through the recipient government leadership and the development partner's proactive involvement.

Resilience coordination is confronted to complexity: integrating developmental evaluation could be useful for strengthening adaptive management in similar future projects

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

## S 022 Turbulent Contexts and Evaluation

### O 067 - Usage of Evaluation Tools to Identify Causes and Trends of Quasi Bellic Conflicts in Latin America (Venezuela, Ecuador and Bolivia)

C. Vela<sup>1</sup>

<sup>1</sup> Ecuadorian Evaluation Network - Self Employed, Pichincha, Quito, Ecuador

**Rationale:** If evaluators seek to be part of the solution on solving conflicts, they need to accept the challenge to evaluate the causes of such conflicts in order to seek a way out of them.

**Objective:** The objective of the paper is to explore how evaluation tools can be of extremely usage for the identification of causes of quasi bellic conflicts in Latin America through evaluating trends of a bundle of countries that express their adoption of a similar ideology, naming Venezuela, Ecuador, Bolivia and Nicaragua.

**Brief Narrative:** Relevance.- the proposed paper is highly relevant because the critical situation of Venezuela, on the verge of being considered a failed State, and the strong crisis with similar trends that are bringing other countries in LAC on the same path. Moreover, it is relevant to Europe because the Latin American Crisis will slash European countries and shares similar causes for political crisis in some Middle East and European countries like Greece.

**Quality and Innovation:** The evaluation has not only applied evaluation tools but, has needed to innovate some of them to adapt the tools to find valid evidences in the era of the non-true manipulation and propaganda. Furthermore, the evaluation goes beyond the business as usual approach of project and program evaluation with identified goals, logic frameworks with defined budgets and evaluating through provided information. Instead, it has worked on trial and fails scenarios for the reconstructions of such logic, and identification of true based on evidences when propaganda sells fabricated realities in order to untwine the political intricacies identified as the root causes of the quasi bellic conflicts. It has been a challenge to identify the final goals of a political project when information is many times is not public. The time frame has also gone beyond the business as usual 2–5 years projects, but extending the time frame to historical bench marks that has marked them.

The evaluation has had a high public interest in civil society movements that feel the effects of the conflicts and are foreseen with anxiety and nervousness the trends they are experiencing. On the other hand, it has also been an opportunity to incentive the culture of evaluation as civil society movements are currently turning their head to understand the importance the evaluation.

The evaluation is highly aligned with the theme of the conference as educating civil society on the evaluation tools and the finding of the present paper has increase their resilience to the current crisis they are living through providing understanding of the trends and hope. The exercise has involved a high number of women of different social classes and education.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 023 Co-Creation and Collaborative Evaluation**

**O 068 - Evaluating Emotion: The Use of Counseling Skills  
in Responsive Qualitative Evaluation**

*J. LaChenaye*<sup>1</sup>

<sup>1</sup> *The University of Alabama at Birmingham, Assistant professor of Educational Psychology and Research, Birmingham, USA*

The level of trust and quality of rapport cultivated between the evaluator and the evaluand directly influence the depth and quality of data collected and resulting overall validity. Unfortunately, qualitative validity often enters under a fog of ambiguity, appearing slightly more concretely in discussions of analysis and presentation yet remaining more abstruse and abstract in other areas of qualitative evaluation. The nuances of the data collection process present a plethora of opportunities for validity successes and weaknesses yet very little structure for best practice is exercised in this realm beyond the focus on proper instrument development and utilization. The purpose of this paper is to present opportunities from helper-centered fields to promote action on qualitative evaluation validity through collaborative analysis and greater stakeholder inclusion, especially in contexts of disenfranchised or culturally diverse communities.

The development of rapport between the evaluator and the evaluand is stressed as vital to best practice, but comprehensive ideals of best practice for practitioners and the relevance of this aspect of qualitative evaluation to the resulting validity of the study is addressed only superficially. In order to strengthen qualitative evaluation practice in this regard, the researchers pose exploration and adoption of rapport-building strategies from participant-centered fields – specifically, the field of counseling. As an area of study primarily concerned with participant well-being and relationship-building, the field of counseling provides a framework of best practice for actively and empathetically engaging with the participant in order to not only provide respectful and professional interaction but also strengthen the relationship between the evaluator and the evaluand.

Adoption of counseling tools is not new to qualitative research; however, the practice has traditionally been relegated to qualitative work involving sensitive topics. The purpose of this paper is to interject these practices more broadly by illustrating ways in which these methods of developing rapport are an integral part of modern responsive and collaborative qualitative evaluation. The presenters intend to expand on their theory-to-application process and its results in evaluation work through examples from their own work assessing participant needs in the affective domain of stakeholder experience. Based on the constructivist tenets of communicating the world as the participant experiences it, the resulting practice encourages effective communication by enriching the experience of the participant in the evaluation process as a necessary and integral step in developing relationships and securing trustworthy spaces for sharing and exploring the topic of interest. This incorporation of counseling best practice as a form of qualitative evaluation best practice further allows us to focus on stakeholder voice, emancipatory and collaborative methods, and transparency as we deconstruct the dichotomy between the evaluator and the evaluand and further strengthen the results of our investigative endeavors.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 023 Co-Creation and Collaborative Evaluation**

**O 069 - Participatory Processes: When the Evaluated Become the Evaluators**

*A. Millard<sup>1</sup>, S. Rauh<sup>1</sup>*

<sup>1</sup> Policy Research Institute, Consultant, Chamoson, Switzerland

Participation has been a concept that has been discussed for quite some time. How participatory approaches are included in evaluation processes varies, however. Indeed, how much participation makes for a participatory process is a question for debate. In this paper, the authors will explore the question of participation more in-depth and look into the experience of including participation in the active analysis of the data. In short, they move away from the participatory process during data collection, and exchange, but towards exploring the opportunities for participation in the analysis.

The authors will explore participation in the analysis by reviewing cases where such approaches have been used (mainly with children), and also share insights of how methodologies have been adapted and implemented to include beneficiary participation in data analysis during an evaluation of gender empowerment programmes. The case study of focus will look specifically at how data was collected, and how programme beneficiaries actively engaged in the analysis of the collected data.

In addition to exploring the actual use of participatory processes in data analysis, and the roles that beneficiaries can play in the analysis of data; the authors will also discuss the logistics and financial implications of this type of participatory approaches.

Overall the objective of the paper is to highlight the opportunities for the participation of beneficiary groups in data analysis processes and the value that including beneficiaries in analysis can bring to the task. The insights that beneficiaries have and their unique perspectives. At the same time, the paper will underscore the investment that needs to be made for such participatory processes to work and be effective.

The authors feel that overall the issue of participation has become so commonplace that there is little discussion on how it can be expanded in the implementation of relatively short routine evaluations. We feel that exploring the implementation of participatory approaches during analysis in the real world is important. Moreover, using the experience of a gender-focused evaluation is particularly relevant. Not least because the intervention under review aimed to increase participatory processes. In this way adopting a highly participatory evaluation process could be seen as further adding to the benefit of the programme itself.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 023 Co-Creation and Collaborative Evaluation**

**O 070 - Once Upon A (Different) Time: The Value of Story  
in a Tempestuous World**

*L. Lee<sup>1</sup>, L. Bremner<sup>2</sup>*

<sup>1</sup> *Proactive Info Services Inc, Vice President, Sandy Hook, Canada*

<sup>2</sup> *Proactive Info Services Inc., President, Sandy Hook, Canada*

In recent years, there has been increasing attention on the value of 'story' within evaluation. We now know that meaning is more easily extracted from 'stories' that resonate with our emotional brain. We all love to listen to stories – and we remember them, or at least recall their essence. When coupled with other data, 'stories' collected through case studies or interviews, or told through photo voice, powerfully communicate the impact of programs on the lives of stakeholders and the community at large. While evaluators must present evidence in support of findings and conclusions, paying attention to the 'story' that underlies the 'data' is critical to communicating impact in ways that resonate meaningfully with various stakeholders. Using stories is an innovative way not only to engage program participants, but also to co-create meaningful – and powerful – evidence of program outcomes. Furthermore, stories resonate across genders and cultures. In many cultures, (for example, Indigenous communities in Canada), stories are an authentic way of conveying knowledge. Stories also provide a vehicle for incorporating shared values of honesty, respectful listening, and reciprocity. However, evaluators should not collect stories in isolation. The most impactful stories are often those told by those intimately involved. This means the role of the evaluator changes from collector of stories to a listener and co-creator of stories. Issues of voice, collaboration, listening, respect and shared meaning become central to the evaluator's toolkit in co-creating impactful evaluative stories. This presentation will provide examples of some co-created stories and the process by which they were created, as well as a discussion of ethical issues that surround story creation and storytelling. Stories give us a way to reflect – and reflect on – different realities that exist in a turbulent, often divisive, and temperamental world. The power of stories to speak truth in memorable ways, especially in tempestuous times, should not be under-estimated.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

## S 024 Evaluation Systems for Public Management

### O 071 - Can Evaluation Save Results-Based Management?

*I. Lindkvist<sup>1</sup>, A. Wilhelmsen<sup>2</sup>*

<sup>1</sup> *Norad, Senior Advisor, Oslo, Norway*

<sup>2</sup> *Norad, Evaluation Department, Oslo, Norway*

Results-based management (RBM) is a governing principle for management of public funds in many countries. The underlying idea is to have a focus on objectives and desired impact, collect information on progress, and use this information to learn and improve. In reality a wealth of research have found that many organisations struggle with results-based management in the public sector. In this paper, we use results-based management in the Norwegian aid administration as a case to demonstrate that 'failures' in RBM, at least in part are due to not seeing evaluation as a key element of RBM. We argue that evaluation and an evaluative culture need to be re-embedded into RBM to allow for resilient and effective management of development assistance and other public funds.

Norway is a generous donor in terms of size of official development assistance and percentage of GNI. Yet, the Norwegian aid administration, consisting of the Norwegian agency for development cooperation (Norad) and the Ministry of Foreign Affairs have been consistently criticised for not being able to document results and for limited use of results-information for learning purposes. This is both surprising as Norway has been branded as an altruistic donor and unfortunate as development assistance is likely to be less effective as a result.

We use existing literature on challenges of results-based management and a recently published evaluation to identify current weaknesses of the Norwegian aid administration's practice of results-based management. We find that the aid administration's response to the critique of not being able to document results appears to have been an increasing focus on indicators and increasing reporting requirements from implementing partners, but with limited use of results information. Ironically, increased use of an indicator-based results-system may not necessarily improve matters as literature on results-based financing tell us that recipients of funding may respond by adapting reporting or programming to maximize funding, rather than effectiveness. We follow Mayne and Rist (2007) and argue that instead what is missing is an evaluative culture and strategic use of evaluation, particularly at portfolio and at the corporate level. This requires that the organisation develops a strategy for monitoring and evaluation. An increased focus on evaluation, may not necessarily increase total monitoring and evaluation costs as results information would be targeted, rather than collecting as much as possible to avoid critique. A move of the aid administration's focus from indicators to development outcomes and management of own portfolios would involve a delegation of authority and responsibility from the administration to partners, i.e. what is currently called trust-based management. Interestingly, this was also the rationale for implementing results-based management in the first place.

In conclusion, results-based management has come under attack in many countries, as a common problem is over focus on reporting and collection of a vast number of data on indicators which are poorly suited for analysis of complex societal challenges. By increasing the use of evaluation, results-based management can create more resilient public management that is better suited to address complex development challenges.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 024 Evaluation Systems for Public Management****O 072 - Time-Driven Activity-Based Costing and Monitoring and Evaluation Systems in the Public Sector. The Case of the General Chemical State Laboratory**A. Vazakidis<sup>1</sup>, E. Kyriakidou<sup>2</sup><sup>1</sup> University of Macedonia, Department of Applied Informatics, Thessaloniki, Greece<sup>2</sup> Independent Authority of Public Revenue of Greece, General Chemical State Laboratory, Thessaloniki, Greece

Activity Based Costing (ABC) and Time-Driven Activity-Based Costing (TD-ABC) are methods for determining true costs. Activity based costing methods identify the activities and assign the cost of each activity to all products/services according to the actual consumption by each. The ABC methods are a recent innovation in cost accounting. In the light of current practices, this paper aims to showcase the need and importance of activity based costing in the public sector, in order to promote more effective financial management, a higher degree of transparency and accountability of public spending and to disseminate a 'cost culture' in the public administration. Additionally, the Monitoring and Evaluation Systems (M&E) is a powerful public management tool that can provide the knowledge capital and support the reforms in public organizations. The control of the costs of the public sector seems more important now, in the context of the financial and fiscal crisis Governments nowadays are rethinking their models, trying to meet the needs of its citizens/consumers of public goods and services. In the case of the Greek General Chemical State Laboratory (GCSL), the TD-ABC method and monitoring system have been applied in order to determine the resources and activities drivers, to allocate the activity cost to the cost objects and to monitor the outcomes. GCSL is a public authority (Ministry of Finance/ General Secretariat of Public Revenue – Independent Authority of Public Revenue) that has some unique characteristics: the highly qualified scientific personnel, the wide range of activity fields and the large number of activities tasks (chemical methods analysis and inspection/audits tasks per activities category). Due to this, the personnel need to be trained in a wide range of activities and at the same time need to specialize (in lab and other technical work and in legal procedures) for the specific activity. In 2013 a structural change occurred across the Greek Ministry of Finance that reduced the number of units (-61%). The consolidation of very small and small units into larger ones has led to a better use and optimal utilization of the personnel. On the other hand, due to a wide range of activities and high specialization, a very high level of knowledge, experience and expertise is required both by the employees and by their supervisor. Due to the international globalized environment, today's public management needs to rely on the principle of the lowest cost in conjunction while maintaining the highest quality.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 024 Evaluation Systems for Public Management****O 073 - Evaluation Methodology for Post-Legislative Scrutiny?***H. Isola-Miettinen*<sup>1</sup><sup>1</sup> National Audit Office, Performance Audit., Helsinki., Finland

Post-legislative scrutiny is part of wider post-legislative doctrine in the area of multidisciplinary legislative science studies, where the evaluation (and audit) has dominant role. As the legislative science concept the post-legislative doctrine refers to the *review of legislation in courts* but also to the *statute review which takes place in parliaments, audit offices, possible in other controlling organs. In this paper I pay attention to the post-legislative scrutiny methods of parliaments, audit offices and other controlling organs (as in UK, EU etc.)*. Post-legislative scrutiny-related tasks are well known activity both in evaluation practice and in performance audit. In my other papers I have studied post-legislative doctrine and the charted various models of post-legislative scrutiny. Post-legislative scrutiny means that the *impacts* of legislation are evaluated *after* the legislation is brought into the force. However, *my argument is that post-legislative scrutiny is not carried out systematically or taken enough seriously as an independent task. It is lacking common core standards*. For example, INTOSAI guidelines are lacking post-legislative scrutiny standards although it mentions the problems of legislation as criteria in performance audit. OECD Regulatory Policy Guidelines (2012) as well EU guides offer some relevant standards. OECD DAC criteria concentrate to program evaluations. *The post-legislative scrutiny as autonomous task means the ex post impact evaluation with purpose to increase the quality of legislation*. It has been stated in relevant literature the problem that there is no fixed model for the post-legislative scrutiny. From the point of evaluation discipline and performance audit practice the topical (common) question is: *Are there found some core minimum method for the systematic post-legislative scrutiny?* The main duty in this paper is to propose the answer to that question on the basis of guidelines, standards and relevant literature. It is known that big problems in post-legislative scrutiny relate to complexity of law, scale of change of legislation, incremental addition of regulations and their interaction. As *Bussmann* states, evaluation of legislation is not following the policy/program evaluation methodology because legislation is complex matter. There exists various kind of legislation, some for ordering the society, some solving the societal problems and reforming society, some are hybrid models (2010). I have studied relevant literature and post-legislative guides (guides for parliaments, performance audit guidelines by INTOSAI etc), EU and OECD Regulatory Policy Guidelines (2012) which possible contribute to the methodology of post-legislative scrutiny.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 025 Reflections on New Designs for Evaluation

## O 074 - “Traditional” Approaches in the Age of RCTs and Big Data

*B. Rohmer*<sup>1</sup>

<sup>1</sup> Oxford Research, Partner, Stockholm, Sweden

**Context:** The trend in evaluation is for more (for lack of a better word) “advanced” experimental and quasi-experimental approaches and new methods that take advantage of technological innovations (such as the possibilities of big data). This reflects both the acceptance of natural-science thinking by evaluators (driven by e.g. high-profile work such as Miguel and Kremer’s 2004 famous “worms” paper) and increasing pressure on policy makers to prove their impact indisputably in the face of turbulent times. In general, this is a good thing, since evaluation’s accountability and learning functions are best served by the most robust evidence available.

However, the fact remains that many interventions are not conducive to these approaches, because of their heterogeneity, limited number of beneficiaries, ethical and logistical barriers to establishing counterfactuals, resource constraints, burdens on stakeholders, complex and evolving theories of change, lack of highly specialised expertise and so on. Such interventions can be evaluated effectively using “traditional” theory-based approaches, but only if good principles are applied to the design and use of specific methods, triangulating between imperfect sources and evaluation management / governance.

These standards are enshrined in guidelines such as the European Commission’s Better Regulation Package and argued for persuasively by evaluators such as John Mayne and Ray Pawson. Nonetheless, based on observations of over 50 evaluations in the European, British and Swedish contexts, the author contends that an (in certain circumstances) unwarranted focus on more fashionable approaches is diverting attention and resources from such principles and thereby reducing evaluation quality.

**Objective:** The paper starts with assumption that “traditional” theory-based evaluation is only likely to produce relevant findings, conclusions and recommendations if the above-mentioned principles are followed. Its hypothesis is that in many cases they are not, partly because commissioners of evaluations (and evaluators themselves) pay too much attention to “advanced” methods that are often unsuitable. The objective of the paper is to determine whether the hypothesis is true, in what contexts and for what reasons, and to propose some solutions to improve the situation.

**Methods:** The paper is based on evidence gathered through (1) relevant literature, particularly evaluation terms of references and reports; (2) the author’s experiences conducting multiple evaluations, mainly on various EU policies and spending programmes, in the fields of public health, customs and tax policies; and (3) a series of discussions with key informants.

**Findings:** The (admittedly anecdotal) evidence largely supports the hypothesis. This leads the paper to conclude that a renewed emphasis is needed on good methodological and management practice for the more “traditional” theory-based approaches that still provide the most applicable framework for a large proportion of evaluations. It also puts forward some ideas for breathing new life into approaches that are familiar and unglamorous but nonetheless vital to evaluation in the real world.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 025 Reflections on New Designs for Evaluation****O 075 - Rethinking Evaluation in US Higher Education: Meeting Evaluation Needs in Rapidly Shifting Climates***K. Winter*<sup>1</sup><sup>1</sup> *Kate Winter Evaluation LLC, Evaluation, Melbourne, USA*

**Rationale:** This paper utilizes the experiences of an independent evaluator of US federally-funded projects in higher education to frame a conceptual argument about the kinds of evaluation approaches needed in today's fast-changing arena. Specifically, the paper addresses the increasing push by US funding agencies for rigid "accountability measures" that assume a one-size-fits-all outcomes measurement system will accurately represent the impacts of programs implemented at diverse institutions. It discusses the shifts in the calls for proposals for funding, many of which now demand studies of causality – regardless of whether such approaches are appropriate. Finally, issues related to implementation failure within some of these projects provide evidence of areas for which attention is needed. Against this background, the paper raises the question of the external evaluator's role, if any, in ensuring projects are implemented with fidelity. The author posits that the external evaluator is ethically bound to offer the evidence needed to push the project leadership towards implementation fidelity, but that it is ultimately the leadership's role and decision how to act on evaluation findings. The paper also explores approaches to framing evaluation designs to encourage valuation of and respect for data and evaluative findings. Various evaluation frameworks are discussed, from Developmental Evaluation (Patton, 2011a) and Utilization Focused Evaluation (Patton, 2011b), to Empowerment Evaluation (Fetterman, Kaffarian, & Wandersman, 2015), Systems Approaches such as Cultural-Historical Activity Theory (Williams & Hummelbrunner, 2011), Realistic Evaluation (Pawson & Tilley, 1997), Principles Focused Evaluation (Patton, 2017), and critical theories such as Feminist Evaluation (Seigart & Brisolara, 2002) and Culturally Responsive Evaluation (Hood, Hopson, & Frier-son, 2005), determining that no one approach best suits the fast-changing evaluation needs of today's US higher education projects, but that a composite of approaches may be required. Attributes of a flexible, composite approach are offered.

**Objectives:** This paper intends to highlight challenges in evaluating US higher education interventions, to put forth potential solutions and paths forward, and to foster dialog with evaluators, funding agencies, and commissioners/users of evaluation regarding where efforts should be focused to ensure successfully implemented interventions that yield evaluable outcomes. The goal is to foster new dialog on appropriate and valuable roles for external evaluators in this unstable political and financial climate.

**Justification:** The paper will likely be of interest and use to evaluators in comparable arenas, project directors in higher education, funding agency representatives, evaluation theorists, and students of evaluation.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 025 Reflections on New Designs for Evaluation**

**O 076 - A Comprehensive Framework to Design and Conduct Evaluation**

*J.A. Ligeró*<sup>1</sup>

<sup>1</sup> *Universidad Complutense de Madrid, Centro Superior de Estudios de Gestión, Madrid, Spain*

The paper presents a theoretical framework through a sequence of 12 steps to design and conduct evaluations. It arises from the need to incorporate an integrative, holistic and not partial theoretical and methodological frame for approaching evaluations with the aim of increasing the use of evaluation.

Although in practice there may be hegemonic or predominant models, theoretically there is not a single way to evaluate. When we talk about evaluation, we are talking about a discipline that has a long historic trajectory with methodological discussions and continuous reflections, which have generated a wide variety of methodological approaches. As proof of this diversity we can just take, as example, the classification developed by Stufflebeam with 22 large evaluative approaches or Marvin Alkin's 41 authors consciously placed on his evaluation tree. In short, there are several approaches, authors and schools that think and work differently. Each can make different sequences, phases or elements in the design and in the way they conduct evaluations.

Following one approach or the other is not innocuous, the choice leads to different products which have different uses and respond to different evaluative purposes or informative needs. Conducting evaluations in a unique manner without considering other alternatives that could fit better to the evaluative needs can be a dissonant exercise, losing usability or, even causing an intentional methodological bias, away from the ethical principles of our discipline.

Therefore, the design of an evaluation should be a conscious exercise about the virtues and limitations of each option, in order to propose the one that better suit the assignment, context and purpose in each specific evaluative work. In addition, this decision-making is not limited only to big choices on methodological approaches, but in many cases also involves elections of lesser extent at different stages of the process. Taking as reference the phases to design and conduct evaluations, for instance, Responsive Clock from Stake, the Rainbow Framework of Better evaluation or the sequences proposed by Alkin among other good references, we propose a framework for designing and conducting evaluations that integrates the main concepts and debates that the historical development of the discipline has been providing. It really is a systematic sequence where the evaluator progresses through 12 key nodes. In each step different issues are deployed, and the answers to these questions allow the evaluator to choose from different alternatives. The final product is a sensitive design, and, therefore, a tailored and specific evaluation. This proposal is not an approach itself but a scheme that allows you to sort the questions and criteria for discerning between the different methodological options and techniques offered by this discipline, allowing rethinking the design according to the criteria of usability and usefulness. The added value of this framework is its holistic orientation towards different theoretical approaches, the adaptation to each assignment and the conceptualizations of the evaluator role as the professional who makes decisions or helps others making decisions about which are the most appropriate options.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 026 Meta-Evaluation and Evidence Summaries**

**O 077 - Is Evaluator-Focused Meta-Evaluation Occurring or Just Theorized?**

*M. Harnar*<sup>1</sup>, *J. Hillman*<sup>1</sup>, *J. Snow*<sup>2</sup>

<sup>1</sup> Western Michigan University, Interdisciplinary PhD in Evaluation, Kalamazoo, USA

<sup>2</sup> Innovated Consulting- LLC, Principal, Washington- DC, USA

Metaevaluation, as a mature concept, has entered an era of application, while evolving past the definition-development stage (Jacob & Affodegon, 2015). This may be more the case for *external, summative metaevaluation*, in which, for instance, evaluation reports are systematically reviewed, yet the program evaluators are distanced from the process. When it comes to evaluators participating in *reflective, formative metaevaluation*, this matured-concept perspective may be presumptive. Our research indicates that evaluators may be less inclined to cooperate in having their work examined; thereby, inadvertently violating key evaluator competencies, standards, and principles. This session presents preliminary findings of evaluator-practitioner efforts towards accountability and improvement. Developing a practice-informed description of *internal formative metaevaluation* that can be used to assess the extent to which the discipline engages in this form of metaevaluation is the aim of this study.

We hope to demystify the concept of internal metaevaluation and bring a more manageable description of practice improvement that provides a useful tool for practitioners, recognizing the very real possibility that evaluators may be hesitant to submit themselves to scrutiny, despite its obvious value.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 026 Meta-Evaluation and Evidence Summaries****O 079 - Evidence Summaries For Policy Making – Using  
A Scoping Review To Inform Implementation Of The SDG14  
Fisheries Targets***G. Carneiro<sup>1</sup>, R. Bisiaux<sup>2</sup>, M.F. Davidson<sup>3</sup>, T. Tómasson<sup>3</sup>*<sup>1</sup> Swedish Agency for Marine and Water Management, Unit for Marine Spatial Planning, Gothenburg, Sweden<sup>2</sup> NIRAS Sweden AB, Evaluation & Results, Stockholm, Sweden<sup>3</sup> United Nations University, Fisheries Training Programme, Reykjavik, Iceland

This paper discusses the experience of producing evidence syntheses for supporting policy making based on a scoping review of the results of development cooperation in the fisheries sector. The adoption in 2015 of the Sustainable Development Goals (SDG) established for the first time a global framework for addressing the most pressing issues affecting the use and health of the oceans. Under SDG14 a new set of targets have been set relating specifically to fishing and the fisheries sector. With this new set of targets in place the global community has pledged greater commitment to ensuring not only the environmental resilience of the resources, but also the socio-economic resilience of fisheries-dependent communities and populations. Sweden has taken a leading position with the co-hosting of the 2017 UN Ocean Conference and is preparing to expand its support to global efforts related to SDG14. However, despite decades of development cooperation in the fisheries sector, no efforts have yet been undertaken to systematically collate and synthesise the experience and knowledge accumulated by the different development agencies and respective partners. The study that this paper is based upon is a scoping review of evaluations of development cooperation interventions in fisheries, the purpose of which is to support the Swedish government in the design and planning of interventions supporting the implementation of the SDG14 fisheries targets. Scoping reviews have been used increasingly in the health sciences and more recently in other domains, including environmental management and development. They are usually undertaken when there are concerns about the feasibility of conducting a full syntheses such as systematic reviews, their main comparative advantages lying in time and cost saving, the ability to address broader topics and include a greater diversity of study designs in the primary literature. In this study we applied a standard scoping review methodology complemented by a quality assessment of the primary literature. In this paper we explore the application of the scoping review method to a novel field of inquiry and analyse its strengths and weaknesses in terms of synthesising evidence, producing new knowledge and supporting policy making processes. We review the key elements and steps of the method and the organisation of the assignment, focusing particularly on the solutions for engaging knowledge users in the review process, and enhancing utilisation of its results. We expect with this paper to contribute to a better understanding of the benefits and challenges of scoping reviews as a methodology for informing complex policy processes dealing with socio-ecological resilience.

NB: I do intend to submit a paper, in spite of the abstract registration system having automatically answered "No" to that question. Please rectify that.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 027 EU Programming and Evaluation

## O 080 - Embedding Ex-Ante Evaluation of New Activities in an EU Agency – Some Lessons Learned

*B. Schmidt<sup>1</sup>, F. Herrera<sup>2</sup>*

<sup>1</sup> Eurofound European Foundation for the Improvement of Living and Work. Conditions, Operational Support Unit, Dublin 18, Ireland

<sup>2</sup> ICF, Research & Evaluation Institute, London, United Kingdom

**Rationale:** EU agencies as European Union bodies are obliged to conduct ex-ante evaluations of any activities 'of significance', by the Framework Financial Regulations of the EU [1].

Implementing these requirements poses a challenge for agencies. It calls for creative approaches to implement this accountability-driven obligation in ways that apply proportionality, ensure that ex-ante evaluations provide added value to the organisation to support the activities being evaluated, as well as demonstrating compliance in the public interest.

As an example, this paper reports on Eurofound's approach to implement ex-ante evaluations to new activities in its current programming cycle from 2017.

#### **Objectives sought:**

1. Provide an account of Eurofound's interpretation of the ex-ante evaluation requirements in the Financial Regulation and their application in practice.
2. Demonstrate how ex-ante evaluation was conducted in practice to new activities of this EU agency during 2017, and share lessons learned from the experience.
3. Contribute to reflections about creative and innovative approaches to ex-ante evaluation in a diversity of organisational contexts.

**Brief narrative and justification (re review criteria):** This paper illustrates the approach to ex-ante evaluations taken by Eurofound, in line with its updated 2017–2020 evaluation policy and programme. This account covers the principles and criteria to be addressed in ex-ante evaluations according to the Financial Regulations, how the agency interprets and implements them with view to new programming elements in its multiannual Programming Document 2017–2020, and how the ex-ante evaluations were conducted and reported on and used in practice.

The paper will present the cases of following new activities which were ex-ante evaluated using this approach:

1. The digital age
2. Monitoring convergence in the EU.

This will be contrasted by an alternative approach to another internally facing activity, to which ex-ante evaluation was applied in a different way.

One way how the ex-ante evaluations were perceived to have been useful in support of the activities was the elaboration of their 'theories of change'. Whilst these were derived from Eurofound's overall programme logic, the application to the specific activities provided valuable and shared insights and reflections into the specific action logics of these activities, which were perceived to be of practical benefit for their continuing implementation.

Finally, the paper will present and discuss some lessons learned from the experience, combining the perceptions from the perspectives of the internal evaluator, the activity coordinators of the activities which were ex-ante evaluated, and the evaluation sponsors (senior management of the agency).

The conclusions will invite broader reflections, discussion and feedback from other evaluation practitioners as well as commissioners, with view to strengthening future ex-ante evaluation approaches in diverse organisational contexts, and thus to more resilient societies.

[1] Article 29.5 of this regulation specifies: *“In order to improve decision-making, the Union body shall undertake both ex ante and ex post evaluations in line with guidance provided by the Commission. Such evaluations shall be applied to all programmes and activities which entail significant spending and evaluation results shall be sent to the management board.”*

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 027 EU Programming and Evaluation

## O 081 - This Paper will Explore how the Programming of EU Agencies' Work and Programmes can be Improved

*Q. Liger*<sup>1</sup>

<sup>1</sup> *Optimty Advisors, EU public policy, Berlin, Germany*

The aim of the paper will be to explore how the programming of EU agencies' work and programmes can be improved in order to ensure that impact evaluations can be performed.

The paper will look at the extent to which counterfactual and theory-based approaches in evaluations can be applied and used in practice in the context of organisation evaluation, and in particular that of EU bodies, interventions but also EU programmes or projects.

The objectives of those evaluations are often to assess the impact of the interventions (or agencies) from a process point of view, assessing the quality of the work and checking whether this has somehow contributed to the desired results, but it's difficult to talk about impact from a rigorous evaluation perspective.

The organisations operate and plan their activities often makes these types of impact evaluations almost impossible. Existing guidelines on evaluation assume that EU bodies assess or benchmark criteria that in fact are difficult to measure.

If clarity is made in terms of what is possible and not possible, agencies could adapt the way in which they plan their activities and ensuring that benefits from an evaluation can really be integrated into the programming cycle.

The paper proposes to draw for the experience of the evaluations done by the EU Agency for Fundamental Rights (FRA) and work developed by EU Agencies in the context of the Performance Development Network. It will be presented jointly by the Head of Sector in charge of planning at the FRA and the Project Manager in charge of the 2<sup>nd</sup> independent evaluation of the FRA.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 027 EU Programming and Evaluation

## O 082 - Gendering Evaluation – Integration of Gender in Evaluation Standards and Alternative Approaches

*A. Wroblewski*<sup>1</sup>

<sup>1</sup> *Institute for Advanced Studies, Higher Education Research, Wien, Austria*

The implementation of the Gender Mainstreaming strategy within the European Union does not only require considering gender in all phases of policy development as well as in all programmes but also considering gender in their evaluation. The German Evaluation Society (DeGEval, [www.degeval.org](http://www.degeval.org)) revised its standards – which are very similar to the Standards of the Joint Committee on Standards for Educational Evaluation – during the last years and passed the new standards in 2016. The Working Group on Gender Mainstreaming (WG GM) within the German Evaluation Society plead for the gendering of evaluation standards and argued that considering the gender dimension in all evaluations – not only evaluation of gender programmes – is a quality criterion for good evaluation. The Working Group has only been partly successful in anchoring the gender dimension in standards. However the forthcoming publication of the new standards contains a chapter which argues that all standards contain a gender dimension and provides examples how to address this gender dimension. The chapter also argues that most evaluation approaches are highly compatible with gendered standards. The Austrian Platform for the evaluation of research and innovation policies (FTEval; [www.fteval.at](http://www.fteval.at)) which is currently revising its standards will formulate an explicit gender standard.

Both approaches have their pros and cons and both require gender competence among all parties involved in evaluation. The WG GM distinguishes between gender competences – which is a must for all evaluators – and gender expertise (see position paper of the WG GM 2015).

In the presentation the gender dimension of standards will be illustrated as well as possible approaches to deal with that in practice. I will focus on selected sub-standards for utility, feasibility, propriety and accuracy. Furthermore the two approaches – gendering of standards versus gender standard – will be contrasted. Finally the relevant gender competence of evaluators will be discussed. The paper will be based on the forthcoming chapter for the publication of DeGEval standards and position papers of the WG GM as well as on experiences of the author in both standards revision processes.

### References:

AK GM [WG GM] (2015), Genderkompetenz von EvaluatorenInnen. Positionspapier des AK Gender Mainstreaming der DeGEval [[https://www.degeval.org/fileadmin/users/Arbeitskreise/AK\\_Gender/Positionspapier\\_Genderkompetenz\\_von\\_EvaluatorenInnen\\_final.pdf](https://www.degeval.org/fileadmin/users/Arbeitskreise/AK_Gender/Positionspapier_Genderkompetenz_von_EvaluatorenInnen_final.pdf)]

Wroblewski, Angela (2018, forthcoming), Gender und die DeGEval Standards, in: DeGEval (eds.), Standards der Evaluation in unterschiedlichen Handlungsfeldern. Einheitliche Qualitätsansprüche trotz heterogener Praxis?, VS Springer.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 028 Theory-Based Evaluation - Innovative Developments****O 083 - Putting the “Theory” Back into Theory-Based Evaluation: Making the Case for Theory-Knitting in Evaluation***S. Lemire<sup>1</sup>, C. Christie<sup>1</sup>*<sup>1</sup> University of California- Los Angeles, Department of Education and Information Studies, Los Angeles, USA

The idea of integrating social science theory in program theories is nothing new in evaluation (Chen and Rossi, 1980; Donaldson, 2007, among others). Yet despite the sustained interest, there is to this day limited empirical knowledge about the extent to which and how social science theory is used in evaluation (Lemire et al., 2018). Moving towards a better empirical foundation, the authors of the present manuscript conducted a systematic review of 107 published theory-based evaluations (and the program theories contained in these). Informed by the review, the present paper first describes the structure and content of the program theories identified in the review. Second, and inspired by Riemer and Bickman (2011), the paper pursues and operationalizes the notion of “theory knitting,” whereby social science theory provides the conceptual foundation for both the program design and corresponding program theory. In conclusion, and towards advancing theory knitting in practice, five illustrative applications of theory knitting are provided.

Theory-based evaluations in general and program theories more specifically are widely used in program evaluation. Towards advancing our development and practical use of program theories in evaluation, the presentation explores how social science may strengthen our development and empirical testing of program theories. The systematic review, on which the proposed presentation is grounded, both builds on and extends beyond past systematic reviews (e.g., Coryn et al., 2011) by including a more comprehensive set of published theory-based evaluations (in the time period from 1990 to 2017); by describing in more detail the components comprising program theories in theory-based evaluation; and by focusing more explicitly on the extent to which and how social science theories are used in theory-based evaluations. Finally, the proposed presentation offers illustrative examples to inspire and promote a stronger integration of social science theory in program theories.

**References:**

- Chen HT and Rossi PH (1980). The multi-goal, theory-driven approach to evaluation: A model linking basic and applied social science. *Social Forces*, 59, 106–122.
- Coryn CLS, Noakes LA, Westine CD, and Schröter DC (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009. *American Journal of Evaluation*, 32(2), 199–226.
- Donaldson SI (2007). *Program theory-driven evaluation science: Strategies and applications*. Mahwah, NJ: Erlbaum.
- Lemire S, Christie CA, and Nielsen SB (2018). Mending the theory gap in evaluation – moving towards theory knitting. In M Palenberg and A Paulson (Eds.) *Evaluation and the pursuit of impact*. New York, NY: Taylor & Francis.
- Riemer M and Bickman L (2011). Using program theory to link social psychology and program evaluation. In MM Mark, SI Donaldson, and B Campbell (Eds.) *Social psychology and evaluation*. New York, NY: Guilford Press.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 028 Theory-Based Evaluation - Innovative Developments**

**O 084 - Like Cat and Dog? Mixing Counterfactual and Contribution Analyses in Evaluating Complex Policies**

*T. Delahais<sup>1</sup>, A. Devaux-Spatarakis<sup>2</sup>, A. Flichy<sup>3</sup>, V. Honoré<sup>3</sup>*

<sup>1</sup> *Quadrant Conseil, Paris, France*

<sup>2</sup> *Quadrant Conseil, Centre Émile Durkheim, Paris-Bordeaux, France*

<sup>3</sup> *Quadrant Conseil, abc, Paris, France*

In the debates over what is impact evaluation, there used to be a conflict in definition between some, defining impact evaluation by its object (impact as the final effect expected from an intervention along a causal chain of expected changes), and others defining impact evaluation by the method used (experimental and quasi-experimental designs, see White, 2009–1). In a way, this debate was loosely opposing theory-based evaluation (TBE) to counterfactual designs. Over the years, this debate has evolved, first with theory-based methods being fitted to answer more rigorously impact-evaluation questions (Stern, 2012); second, because evaluators have been experimenting more and more mixing different methods in their practice.

Authors have explained how using a Theory of Change could really improve a counterfactual evaluation (White, 2009, Funnel, 2011) and in a way, benefit from “the best of two worlds” (Van der Knaap, 2008). In such settings, counterfactual is the primary source of causal inference, and TBE is used to guide “qualitative” inquiries as to the how and why does the intervention work or not.

However, there are many situations in which impact evaluation is primarily done through a TBE. TBE is open to many designs. In Contribution analysis specifically, which is our design of choice, three questions are asked: what happened? What is the evidence of a contribution of the evaluated intervention? What is the evidence of a contribution of other interventions? A counterfactual would naturally fit in answer to the 2<sup>nd</sup> and 3<sup>rd</sup> questions. But what happens in practice when counterfactual approaches are used within a theory-based evaluation, and especially a contribution analysis?

To answer, we use three recent examples of TBEs relying at least partly on counterfactual approaches as part of their general design: the evaluation of the effect of European Structural Funds on Employment in a French region; the evaluation of an Investment programme supporting industrial innovation at French level; and the evaluation of the Youth Employment initiative in France. These examples are used to consider how the two approaches to causality actually mix; how this mix is implemented; and what is its actual value added in light of the results they bring, their credibility and their use for decision makers.

We find that in our experience, 1/ differences in paradigms, expertise, timeline, and data needs all tend to make TBE and counterfactual parallel approaches rather than a single evaluation, each one with its own merits and drawbacks; 2/ that the counterfactual approach (including data collection requirements) helps increasing the robustness and credibility of the overall evaluation, even when it is not possible to perform it (because it was tried at least); and 3/ on effects that are not considered by the counterfactual, upstream and downstream the causal chain.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 028 Theory-Based Evaluation - Innovative Developments

## O 085 - Innovations in Theory-Based Evaluation: Using Nested Actor-Based Theories of Change to Evaluation a Complex International Development Intervention

A. Koleros<sup>1</sup>

<sup>1</sup> Palladium, Research- Monitoring and Evaluation, London, United Kingdom

With origins in the field of evaluation, over time development practitioners have come to see a number of different uses for theories of change, including as an approach to program design and program planning, as a tool for program monitoring and accountability, and as a mechanism for communication (Vogel, 2012). Theories of change have now become common in evaluation, although what comprises a theory of change varies widely, as does their actual use in evaluation.

At the same time, interventions have become more complex, making the development and use of theories of change more challenging. This shift is driven in large part by a growing recognition from theorists of the need to better engage with concepts of complexity (Ramalingam, 2013).

This session explores one approach to tackling some of this complexity through the use of nested actor-based theories of change on the impact evaluation of a complex police reform programme in the DRC. We describe how we used the programme's original theory of change to build layers of nested actor-based theories of change to design the evaluation. We then discuss the type of evaluation findings this approach was able to generate.

This is followed by a discussion on the advantages of using this approach compared to other theory-based evaluation approaches. First, we found that the focus on actors was a useful way of understanding causality in a complicated programme, with multiple component activities working with multiple target populations to bring about long-term and sustained change in the population at large.

We also found that our ability to engage programme stakeholders into the actor-based approach was a significant advantage. We found that stakeholders were better able to articulate and reflect on changes when put in the context of specific actor groups as opposed to more theoretical causal impact pathways that are often used in theory-based approaches.

Finally, an actor-based approach helped us to usefully build in an assessment of the likelihood of sustainability of the intervention. By focusing on how actors would need to sustain a behavioral change, we could assess factors that would need to remain in order to create these favourable conditions and better determine which were likely to remain in the absence of programme funds.

We conclude with a reflection on how these concepts can be better integrated into theories of change in order to improve the robustness of both intervention and evaluation design.

Ramalingam, B. (2013). *Aid on the edge of chaos: Rethinking international cooperation in a complex world*. Oxford: Oxford University Press.

Vogel, I. (2012). *Review of the use of 'theory of change' in international development*. Department for International Development. Retrieved from [http://www.oxfamblogs.org/wp-content/uploads/DFID-ToC-Review\\_VogelV4.pdf](http://www.oxfamblogs.org/wp-content/uploads/DFID-ToC-Review_VogelV4.pdf)

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 029 Evaluation and The Refugee Crisis**

**O 086 - Migration and Educational Outcomes: Assessing the Impacts of Refugee Status on Learning in Uganda**

*M. Nyanzi*<sup>1</sup>

<sup>1</sup> *Twaweza East Africa, Data and Voice, Kampala, Uganda*

About half of the sixty-five million estimated international migrants are children of school-going age. Tackling the education-related challenges such children and their families face is a real priority for international development today, or else achieving quality education for all might remain a distant dream.

In this cross-sectional study, we analyze data from the recently-concluded Uwezo baseline assessment of basic competences in four districts hosting forcibly displaced persons in Uganda. Using a number of variables to proxy for refugees' levels of integration in the host communities, we find associational evidence of a significant role played by local integration on the education outcomes of refugee children – including school attendance and proficiency in basic reading and counting.

These findings bring into critical focus, at a global stage, the importance of host country national and sub-national policies and practices on such integration aspects as language of instruction in school, and refugees' involvement in local economic and other social activities in creating resilient societies. The evaluation of education in crisis situations like migration presents multiple areas of evaluation like the level of integration to provide evidence aimed at enabling a better understanding of what it will take to improve education outcomes.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 029 Evaluation and The Refugee Crisis

## O 087 - IKEA Initiatives to address the Refugee and IDP Crisis TWG on Private Sector Evaluation

*N. Mathieu*<sup>1</sup>

<sup>1</sup> Former World Bank and EBRD staff member, Evaluation, Washington, USA

IKEA is currently intervening in two ways: (a) in partnership with UNHCR and with the collaboration of the company Better Shelter, IKEA produces and sells solid frame shelters to equip refugee and IDP camps; (b) in partnership with the Jordan River Foundation, IKEA employs refugee and IDP artisan women to weave textiles products for sale through its network.

These interventions bypass the usual donation structure under CSR where a company transfers money to a NGO or a charitable association that implements a program. IKEA instead directly associates partners and ultimate beneficiaries to its company commercial objectives while fulfilling a common objective. Both public and private interests converge towards a humanitarian goal. The two projects are being implemented under private sector budget discipline and commercial standards of product design and delivery.

The sustainability of both IKEA operations depends upon the continuous feedback from beneficiaries, either as consumers or as employees. The capacity to correct from deviations of company objectives under market constraints should help consolidate the sustainability. In this framework the gender equity principle should be substantially met since women and men benefit equally of the services provided by the shelter in the first project and woman are can further develop their already acquired weaving skills in the second project.

The projects are at early stages of implementation. The shelter project started in 2013 with new technology prototype shelters being tested, and three years later 6,000 shelters have been sold. In 2017 10,000 shelters were placed on hold, as they did not fully meet the fireproof requirements from Switzerland and Germany. These shelters are to be deployed only in refugee camps where the space arrangements are currently adequate for protection from fire hazards. Meanwhile IKEA in collaboration with Better Shelter keeps improving the fireproof quality of the shelter.

The handicraft project is even more recent. Benefitting from similar past experiences of integrating handicraft work in limited textile collections to be sold commercially, IKEA in partnership with the Jordan Foundation started in 2017 a new collection project of textiles to be woven by employing refugee and IDP artisan women. 100 artisans were employed in the starting year with a plan to move relatively quickly to 400.

Within this short time of implementation a full ex-post evaluation is still remote. Even a real time evaluation could be premature given the limited information available. The exploratory approach suggested here consists in making a preliminary assessment of the project design, the short term impact and the remaining potential using the basic DAC criteria. This would help to find under what conditions direct private sector initiatives could be a good addition, or alternative, to more traditional CSR donations to charitable organisations or NGOs, while addressing the refugee and IDP crisis.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 029 Evaluation and The Refugee Crisis

## O 088 - Challenges to Evaluating Policies and Initiatives Addressing the Migration/Refugee Crisis

*A. Heetman*<sup>1</sup>

<sup>1</sup> *Optimity advisors, Government - EU, London, United Kingdom*

Since 2015, Europe has seen an unprecedented migratory flow into the EU. In 2015 alone, Member states reported over 1.8 million irregular arrivals of migrants and refugees. The EU and its Member States have been adopting a range of different approaches to address the refugee crisis, such as restrictive border management policies, agreements with key countries of transit (e.g. EU-Turkey agreement) and preventative initiatives such as information campaigns, as well as through the criminal justice system (arresting and prosecuting smugglers etc.).

Over three years have passed since the height of the refugee crisis and the adoption of the European Agenda of Migration by the EU. However, the question remains of how effective these policy measures have been, especially when all are taken together. Some policies seem to intrinsically contradict each other, such as policies relating to the right to asylum vs. those promoting restrictive border management. In addition, many have argued that closing one migratory route will only open another more dangerous route into Europe. So far, evaluations of these migration related policies have been done in a relatively limited way, often with differing results.

Proper evaluation requires the establishment of an intervention logic. However, it seems that in the field of migration, the very starting point of the intervention logic, the objectives, is not always clear: for e.g. is the aim to reduce the number of migrants arriving in the EU or is the aim to ensure the safety & health of migrants? Another challenge is the issue of the 'dark number': a decrease in detections of irregular migrants could be an indicator of an effective measure (decrease in irregular entries), but equally could point to an ineffective measure (inability to detect all irregular entries) etc.

This paper will try to better understand the difficulties in undertaking such evaluations, and to propose some opportunities for the evaluation field in this area. This paper will cover the following questions:

- What are the latest attempts of evaluating the effectiveness of the different migration related policies and measures at EU and national level?
- What data can be used to measure effectiveness and what indicators are already being used?
- What are the key challenges of doing evaluation in this field?
- Opportunities: What other data could be collected and used? What innovative evaluation methods and approaches have been used that seem to work?

The paper is based on different pieces of research carried out by Optimity Advisors in the past five years in the field of migration and home affairs for the EU Institutions, such as the Evaluation of the practical application of the Return Directive (2008/115/EC), the ex-post evaluation of European Border Fund (EBF), the study on the smuggling of migrants and the study on communication channels used by west African migrants to obtain information in countries of origin and transit, with particular focus on online and social media.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 029 Evaluation and The Refugee Crisis****O 089 - Evaluating Social Cohesion to Understand Resilience Among Internally Displaced People Living in Somalia's Urban Areas**S. Griffiths<sup>1</sup><sup>1</sup> Coffey International, Evaluation and Research Practice Leader, London, United Kingdom

In the first half of 2017 Somalia experienced the worst drought for decades. The Humanitarian Information Unit (HIU) estimates (2017) that 1.1 million internally displaced people currently live in Somalia, and at least 548,000 additional people have been displaced since November 2016 due to the drought. Most people displaced by the drought have settled in urban areas, particularly Mogadishu and Baidoa in South West Somalia. This situation has been exacerbated by the return of about 56,000 former Somali refugees from Kenya's Dadaab refugee camp to Somalia. In Somalia, displacement is an urban issue. Somalia's main cities are a pull factor to those who have been forced to leave their homes, or who are unable to return to their homes. They have greater access to basic services and in many cases are unlikely to return home. Donors, NGOs and implementing partners providing humanitarian support recognise the importance of resilience as a bridge allowing displaced people to smoothly transition from cash assistance to more sustainable livelihoods. However, displaced people living in Somalia's urban centres are confronted by a range of complex and dynamic contextual factors. This complexity is compounded by under-resourced municipal or regional administrations who now face a massive disruption to the delicate equilibrium that they previously managed through consultation and negotiation with clans and other constituent parts of their urban communities. In these contexts, developing the urban resilience of communities that have suddenly been brought together by this catastrophic shock remains a significant challenge for local authorities, donors and NGOs.

Simon Griffiths (Coffey International), Linda Beyer (Visiting Scholar, Urbanisation and Well-Being Unit, African Population and Health Research Centre) and Dr Katharine Downie (Head of Data Quality Assurance, M&E and Knowledge Management – Somalia Resilience Program) are developing an evaluation approach that provides key decision-makers with data and analysis that could be useful in managing the distribution of scarce resources as equitably as possible to such diverse urban communities. Our paper and presentation builds on current work by organisations such as: the Somalia Inter-Agency Standing Committee Framework for Durable Solutions for Internally Displaced Persons; ReDSS; and the Food Security Information Network's 20 member Resilience Measurement Technical Working Group. A key emphasis of our approach is a greater focus on social cohesion at the centre of an evaluation framework that also measures resilience capacities, contextual factors, household characteristics, well-being outcomes and shocks and stresses. Planning and enacting urban resilience in Somalia is inevitably a contested process that local authorities, NGOs and donors need to manage. It requires trade-offs to achieve the level of social cohesion needed for inter-connected urban systems to function properly. These trade-offs should be informed by robust evidence of the social and economic dynamics that drive social cohesion among diverse populations brought together by Somalia's worst crisis in decades. Our objective is to demonstrate the need for a greater focus on social cohesion when evaluating the resilience of internally displaced people living in diverse urban communities and explore opportunities for further development of the approach that we are proposing.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 030 Auditing and Regulatory Impact Assessments

## O 091 - Evaluating Regulatory Impact Assessment in Switzerland

*C. Morier*<sup>1</sup>

<sup>1</sup> *Performance and Evaluation Competence Center, Swiss Federal Audit office, Bern, Switzerland*

Do Regulatory Impact Assessment (RIA) ensure that legislative proposals address the identified problems in an adequate manner and that the solutions proposed will not have unanticipated side effects on parts of society? Following OECD recommendations, numerous countries prescribe an RIA at the early stage of the policy process for the formulation of new regulatory proposals. Although it is a difficult exercise, based on predictions to the future, government agencies must analyse the different options and their possible impacts in order to select the most appropriate measures and to provide the public and parliaments with dependable, transparent information.

Ex-post evaluation can in turn add value by examining if governments actually implement RIAs, and how this contributes to improving regulation. The Swiss Federal Audit Office (SFAO) conducted such an evaluation in 2016. The key driver for commissioning this evaluation was a high-profile example of a corporate tax reform, where the Government dispatch had grossly underestimated the future loss in tax receipts. In total, the SFAO examined 50 dispatches of the government. The goal was to assess how the federal agencies assess the consequences of legislative projects (respect of the guidelines, transparency and reliability).

The SFAO combined different methods: content analysis, survey, qualitative-comparative analysis (QCA) and case studies. The content analysis was best suited to gaining an overview of the work of the administration. The SFAO developed minimal standards and assessed if the dispatches reached them and were of sufficient quality. This methodology proved to be very fruitful although the results were quite disappointing regarding the agencies assessment practices. The use of some other methods happened to be challenging, especially the case studies. The purpose was to verify years afterwards if the prospective impacts had been correctly anticipated, which was not an easy task.

The evaluation report received great media attention. A recommendation by the SFAO for implementing an external quality control, also proposed by a parliamentary motion, is still being debated within the administration. Ministers and agencies fear a prolongation of the legislative process that is already long in comparison with other countries. Above all, it is likely that they want to avoid a loss of power that would necessarily result from the implementation of an external control of the conducted impact assessments.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 030 Auditing and Regulatory Impact Assessments****O 092 - Extending Evaluation Towards Regulatory Impact Assessment to Reduce Regulatory Burdens: A Neo-institutional Study of Context, International Impulses, and Finnish Experiences**P. Ahonen<sup>1</sup><sup>1</sup> *University of Helsinki, Professor, Helsinki, Finland*

This paper examines the extension of evaluation comprised of regulatory impact assessment (RIA) of a specific type with special reference to the country case of Finland. The paper is focused upon the contextual conditions of applying this specific RIA type, the relevant international impulses, and the Finnish implementation of the RIA type. This type of RIA should enable the government to put into practice the international "one-in, one-out" principle to freeze the costs of regulation to business companies. Testing three neo-institutional hypotheses gave the following results. First, using notions introduced by J.G March and J.P. Olsen, in the contextual conditions concerning the specific RIA type a conventional "logic of appropriateness" rather than a rationalistic "logic of consequence" prevails, as indicated by such practices as the flexible rather than strict implementation of this type of RIA. Second, besides a formal rationality ensuring the simultaneous observation of ends, means and side effects, legitimation-seeking emulation of international examples with special reference to those of the United Kingdom, Canada, and Germany the specific RIA also underlies this RIA type. Third, institutional rationalized myths play roles in implementing the examined type of RIA with exaggerated national beliefs in the validity, reliability and objectivity of the institutional models and institutional scripts used to establish this type of RIA.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 031 Evaluating Youth Employment Programs

## O 093 - Catalyzing Digital Employment Opportunities and Skills Training for African Youth: Assessing Impact Using a Case Study Approach

*C. Smit<sup>1</sup>, E. Massey<sup>1</sup>, M. Jakoet<sup>1</sup>*

<sup>1</sup> *Genesis Analytics, Evaluation for Development, Johannesburg, South Africa*

Job opportunities, particularly among the youth, are not keeping up with the rapid growth in Africa's populations. Youth unemployment is increasingly a challenge in many African countries and overcoming this has come to the forefront of policy and donor agendas as a crucial means of benefiting from the demographic dividend and building more resilient societies.

One global Foundation launched an initiative in 2010 to catalyse new, sustainable employment opportunities and skills training for African youth, with a focus on the ICT sector. In 2016, Genesis Analytics became the ongoing monitoring, evaluation and learning (MEL) partner to the initiative. As part of this partnership, Genesis conducted an evaluation of the initiative from 2010 to 2016. The purpose of the evaluation was to assess the impact of demand driven training and job placement in ICT-related sectors on youth reached and their circle of influence, and to unpack the factors that influence this impact.

This presentation aims to provide an overview of this initiative, the role of MEL partner and how a case study approach was selected to be the most effective means of assessing impact and generating lessons to inform programmes targeting young people.

The evaluation was guided by seven overarching evaluation questions and focused on South Africa, Ghana and Kenya. As a starting point, it was important to understand the context of youth unemployment and the huge number of factors (qualifications, labour market conditions, proactiveness, networks, to name a few) that influence job creation and impact on young people and their sphere of influence. Because of all the factors that influence an individual's employability and subsequent impact, the evaluation team decided that it was important to explore each individual's impact in depth rather than a narrow perspective of a larger array of young people. As a result, a case study approach was chosen. This meant that there was an emphasis on collecting qualitative information to build a comprehensive understanding of impact per individual and differences in individuals' vulnerability; and to then look across cases to unpack the drivers of impact. The evaluation team also conducted focus group discussions with similar groups of young people in all three countries to assist in locating case study findings within a broader landscape and to assess the extent to which views were shared by others.

The evaluation generated a range of helpful findings related to how the lives of the young people that benefited from the initiative were changed, how the livelihoods of their sphere of influence were impacted and the initiative addressed the major barriers that prevent young people from finding jobs.

The evaluation also generated a number of helpful lessons for conducting ex-post evaluations of youth-focused initiatives within a limited budget, namely:

- Less traditional approaches for recruiting FGD participants were more effective e.g. WhatsApp, and could be leveraged for follow up fact checking and verification
- Criteria used to select cases needs to be rooted in what is realistic and feasible
- Peer reviewers are essential for validating and sense-checking a chosen approach and selection criteria

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 031 Evaluating Youth Employment Programs

## O 094 - Evaluating Interventions Targeting Young NEETs in Romania: Data Collection Challenges and Results

*L. Trofin*<sup>1</sup>

<sup>1</sup> Qures Quality Research and Support, External evaluator for Ministry of European Funds, Bucharest, Romania

**Rationale:** The presentation aims to share the experience gained during the evaluation carried out in Romania in 2016: “Assessment of the Human Capital Operational Programme contribution to the employment increase among the young NEETs” (Not in Employment, Education of Training). In 2011 only 34% of Europeans aged between 15 and 29 were employed. In response to this problem, the European Commission developed the EU Youth Strategy for the 2010 – 2018 period, and concrete instruments such as the 2012–2013 Youth Opportunities Initiative, the Youth Guarantee and the Youth Employment Initiative. Despite an overall progress registered years later, youth unemployment remains high in many Mediterranean Member States. In Romania in 4 out of 8 NUTS II regions youth unemployment was higher than 25% in 2012. In 2016, only 67.702 young NEETs were registered with PES (Public Employment Service), out of approximately 400.000 estimated. Considering the particularities of this target group, both in terms of reach-out and effects to be obtained, i.e. increased employability and employment, particular data collection methods and approaches are needed for rigorous evaluations to be carried out.

**Objectives:** Our objectives are two-fold: on one side we intend to share our experience regarding (1) data collection from young NEETs in Romania, which posed specific challenges, (2) the experimental application of the Prospective Evaluation Synthesis (PES), a theory-based type of evaluation (as per latest edition of EVALSED) and to discuss the results of the evaluation as such with a view to inspire policy change.

**Brief narrative and justification:** The evaluation was theory-based, and data was collected through interviews, 1 survey among young NEETs benefiting of EU-funded support, focus groups and case studies based on 2 country-wide projects. Hypotheses was set up based on Prospective Evaluation Synthesis (see EVALSED Sourcebook page 60) which took into account evaluation reports of similar interventions provided by the beneficiary, and analysis of European and national legislative and policy context. Survey was challenging to apply as EU projects were closed and the contact data, if existing, was not available anymore. Response rate increased when applied an approach closer to „generation characteristics”, i.e. survey was sent by texts and could be filled out on mobile phone. Data collected was triangulated and hypotheses validated, where the case. Given the relative novelty of the subject, we do expect to contribute to improve evaluators' capabilities, particularly under criterion 1.3 Understands the potential and limits of evaluation instruments and tools and 2.1 Demonstrates capacity to manage and deliver evaluations. Simultaneously, considering the new methods used (i.e. Prospective Evaluation Synthesis) and innovative ways to survey young NEETs (which went beyond “business as usual”), we expect to generate fruitful debates and provide the participants with important takeaways.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 032 User-Driven Evaluation

## O 096 - Evaluators, Empathy and Policymaking – Can a Better Understanding of the Policymaking Process Enhance the Use of Evidence?

*M. Van Parijs*<sup>1</sup>

<sup>1</sup> *Enabel - Belgian Development Agency, Independent Evaluation, Brussels, Belgium*

With global issues at stake and an increased demand of citizens for effective policy outcomes, 'evidence-informed policymaking' has entered in the spotlights. There appears to be little doubt that evidence is useful for policy decision making, it remains unclear however how this needs to be done. *Evidence-use* in policymaking is often presented in a simplified and linear way as a matter of generating- accessing-using evidence, whereas real-world policymaking is more complex. If the evaluation community aims for evaluation evidence being used in the policymaking process, it is needed to gain understanding of real-world policymaking and of how evidence interferes in the policymaking process. Increased insights on how evidence enters, behaves and is used in the policy arena, is expected to lead to a better understanding of the policymakers' needs, and consequently of the role of evaluation community, beyond making evidence accessible and communicating evidence.

The presentation will highlight conclusions of a literature review that started with the question how the gap between 'evidence is accessible' and 'evidence is used' could be bridged. The presentation challenges the mainstream perception of communication as a way to 'bridge the gap'. The main line is that evidence is more political than generally acknowledged. Political factors can induce different forms of bias in the creation, selection and interpretation of evidence; whether evidence is used in the policy arena or not is not a mere matter of 'robustness'. Recognising the 'meaning-making' aspect of the policymaking process, clarifies why evidence does not function on its own. It is thus necessary for the evaluation community and especially evaluation units to take into account the need for interaction on evaluation evidence, right from the design phase of evaluation onwards, and to recognise the politics of evidence. Evaluation units are also to move towards the role of 'evidence translator' – helping with interpreting the value of evidence for policymaking, as well as with prioritizing and negotiating the need for 'new' evidence. The presentation builds upon insights from the analysis of the use of evidence policymaking and of evidence and argues that evaluations, focusing on drawing lessons for policymaking, are to be considered as a 'service' to policymaking, and therefore, have to be adapted to the nature of policymaking.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 032 User-Driven Evaluation**

**O 097 - User-Driven Evaluation and Innovation in Recovery Oriented Services in the Welfare and Health Sector**

*K.J. Johansen*<sup>1</sup>

<sup>1</sup> Norwegian University of Science and Technology, Medicine and Health Sciences, Trondheim, Norway

In Norway it has been mandatory for a few years to include service users in health care research. This has challenged researchers to find good ways to fulfill this requirement that give real impact on the research process and implementation of results. It has also resulted in a high demand for users who are interested in participating in various projects and who have the knowledge and skills to become real participants and not just hostages to meet the government's orders. In recent years this has been brought even a step forward by initiating user-driven evaluation and development processes. This has especially been initiated regarding recovery services in the Welfare and Health sector. It is a big challenge to accomplish that service-users evolve from being recipients to becoming empowered service evaluators and transformers. In Norway this has been carried out through cooperation with some municipalities, hospitals, universities and user-controlled centers in Mid-Norway. There has been tested out different approaches to see what is working. Enabling the implementation of such processes has been a matter of several dimensions; like involving the right partners, initiating empowerment processes, give necessary training and motivation and systematically facilitate and follow-up users that are involved. In the presentation is offered information about the basic thinking and overall approach and reflections about how this collaboration has functioned. It will also be given insight into practical implementation steps involved and the concrete actions and partnerships.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 032 User-Driven Evaluation****O 098 - Not About Us Without Us – How Far Do We Go?***B. Larsson<sup>1</sup>, W. Doughty<sup>2</sup>*<sup>1</sup> *BIMLarsson & Associates, Principal, Edmonton- AB, Canada*<sup>2</sup> *University of Alberta, Student Success, Edmonton, Canada*

Recent literature on evidence use in public policy argues that bringing credible and rigorous evidence to decision makers is not sufficient; the evidence needs to be brokered (Olejniczak, 2017). Michael Patton has argued, over many years, that inclusive and participatory evaluation leads to increased use of the evaluation results, strengthened capacity, and more ownership of what has been created through the evaluation (Patton, 2008). Furthermore, the Canadian Evaluation Society's Competencies for Credentialing requires that the evaluators must "respect all stakeholders and strive to build trusting relationships." Mitroff's (1980) seminal work in identifying stakeholders and recognition of the value of stakeholder engagement continues to be supported in the literature (Bourne, 2008; Sangle, 2010; Manetti, 2001).

As evaluations continue to increase in complexity, diversity and risk the need to be inclusive of stakeholders persist. The presenters have, throughout their own practice seen, how establishing a sound relationship is a critical component in human services evaluations. Trusting and respectful partnerships create leverage and buy-in that enhance the depth and understanding of the phenomenon under assessment. It also creates a working environment that moves the evaluation forward and allows it to be better situated and more responsive to project challenges such as restricted timelines and evaluation expectations. The evaluator has to balance the need for objectivity and create a true partnership. Without question, the evaluation must rest on the fundamental pillars of sound methodology and rigorous approaches. How does the evaluator then resolve concerns that stakeholders may have about methodology, approaches, etc. How far does this the engagement of stakeholders extend? When is it time for the evaluator to step away and allow the evaluand to navigate and respond to the evaluation findings? The presenters will share the foundation and context for relationship building between the evaluation team and stakeholders, with a specific focus on the role of the project team. They will then proceed to look at the value of extending this relationship beyond interpretation of findings to implementation of recommendations. What does it truly mean to "embed" an evaluation? This will be accomplished by providing highlights from the literature, sharing a few case examples, and facilitating a conversation about "how far is too far".

Wednesday, 3 October 2018  
14:30–16:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 033 Turbulent Contexts and Evaluation****O 099 - Evaluations in Emergency, Recovery and Resilience  
Context: Lessons from Myanmar, The Philippines, and  
South Sudan***H.J. Garcia<sup>1</sup>, J. Pinat<sup>2</sup>, M. Igarashi<sup>1</sup>, L. Belli<sup>1</sup>, M. Bruno<sup>1</sup>*<sup>1</sup> Food and Agriculture Organization of the United Nations, Office of Evaluation OED, Rome, Italy<sup>2</sup> Food and Agriculture Organization of the United Nations, Regional Office for Near East and North Africa, Cairo, Egypt

In accordance with its mandate, the Office of Evaluation of the Food and Agriculture Organization of the United Nations (FAO) conducts evaluations of emergency and recovery interventions. These evaluations aim to ensure accountability and to collect lessons for future responses and corporate learning. This paper presents the highlights of three evaluations conducted in Myanmar, the Philippines and South Sudan, including the approaches, tools, challenges, results and lessons learned drawn from these evaluations. Based on these findings, the document suggests practical recommendations for evaluators and evaluands. Regarding Myanmar, the paper reviews evaluations of recovery projects conducted after Typhoon Nargis in 2008, as well as the emergency projects in Rakhine State that responded to a combination of typhoons, flooding and communal violence since 2010. The interventions used to bridge emergency and recovery activities are discussed, such as Village Fishery Societies. The paper also reviews two emergency evaluations: i) the Level 3 response conducted in an unstable political environment in South Sudan, including where the programme is situated in a deepening humanitarian crisis, and how an impact assessment was implemented; and ii) FAO's Level 3 response to Typhoon Haiyan in the Philippines in 2013. This paper also highlights how monitoring and evaluation (M&E) and Accountability to Affected Populations was implemented.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 033 Turbulent Contexts and Evaluation

## O 100 - What Are We Evaluating For – Better Delivered Programmes Or Increasing Research Bank?

*P. Patel*<sup>1</sup>, *E. Daniel*<sup>2</sup>, *R. Narayanan*<sup>3</sup>

<sup>1</sup> *Catalyst Management Services, Technology for development, Bangalore, India*

<sup>2</sup> *NA, Na, Mumbai, India*

<sup>3</sup> *Catalyst Management Services, Results and Insights, Bangalore, India*

Monitoring and evaluations are mandated by all funders for the programmes supported by them. However, the objective of these is often restricted to transparency or return on investment or a formality. In larger programmes with higher stakes, the agenda of contributing to larger repository of knowledge takes over. The key objective is then to ensure that new and unknown evidence is generated which the future generation of implementation designers can account for. However, in all of these processes, the key stakeholder, the implementing agency, is often missed.

Programmes and the communities are seen as repositories of data and information which are used across multiple years and after the programme has ended to answer what worked in the programme, what has not and what would an ideal programme look like. This raises an ethical dilemma for the evaluator – who is the end consumer for his research? While we at CMS understand the need and importance of generating evidence for the larger audience, we give greater importance to the short term audience – the implementer.

The framework for implementation of CMS is driven the PALA framework – Performance, Accountability, Learning and Action. The programme is at the centre of this framework and monitoring and evaluation works towards ensuring that it has access to the necessary data and results to perform all the above mentioned roles.

Another interesting dimension to evaluations is the use of innovative technology methods with IT solutions to collect data, provide immediate feedback to the programme through smaller subsets while generating larger repositories for evidence generation. This helps the programme to understand the key gaps in its current implementation, do course correction and also give a pathway for improvement. As this process is data driven, it disengages itself from the dilemma of working for the present vs working for the future.

The third phase of Avahan India AIDS Initiative, implemented by a consortium of CSOs called Swasti and Vrutti is an example of such an intervention. The team supported by CMS, designed, developed and implemented an IT solution called Taaras. This solution tracked the progress of the programme on real time against the key outputs and outcomes. This helped the implementation team to prioritize the population to be reached, bring in geographical focus and also to ascertain if it is moving in the right direction. This helped generate a cohort database which was then used for research and dissemination for external stakeholders. This was supplemented by annual sampled outcome monitoring studies which while confirming the data from the monitoring sources also provided additional dimension to the cohort data.

Thus, to conclude, in the current generation of research and implementation programmes, equal importance needs to be given to generating evidence for immediate consumption and for long term consumption. With this outcome in foresight, innovations needs to be integrated to all elements of evaluations.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 034 Integrating Gender Perspective in Evaluation**

**O 102 - Integrating a Gender Perspective in Country Development Strategies and Programs: Lessons from the World Bank Group's Country Program Evaluation in Rwanda**

*Y. Hirano*<sup>1</sup>

<sup>1</sup> *The World Bank Group, Independent Evaluation Group, Washington- DC, USA*

Integrating gender perspectives in country development strategies and programs has emerged as a critical concern in the international community. Since the United Nations adopted the first resolution on gender mainstreaming to guide the implementation of global commitments on gender equality and the empowerment of women in 1997, both international organizations and developing countries have taken significant measures to integrate gender perspectives in their policies and programs. Consequently, there has been an increased production of guidelines and tools on how to integrate gender perspectives in evaluation. However, there has been little guidance on how to address key challenges such as the lack of available information and other constraints in assessing gender mainstreaming in practice (WBG, 2016). In addition, assessing the gender dimension in the achievements of country strategies and results of projects has been a challenge, as it is often the case that gender mainstreaming is only measured at project entry.

The focus of this paper is twofold. First, this paper introduces the approach and methods that the World Bank Group (WBG) has employed to address gender issues in evaluation. Second, the paper introduces a case study based on the WBG's Country Program Evaluation FY09–17 in Rwanda, which applies the approach and methods for assessing gender mainstreaming. For this particular case, the presentation addresses how well the WBG has been integrating gender issues in its country strategies, investment projects and knowledge and advisory work, the extent to which gender-related objectives have been achieved, the potential contribution of gender-related issues to shared prosperity, and finally the overall lessons learned with regard to the WBG's operational experience in gender mainstreaming.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 034 Integrating Gender Perspective in Evaluation**

**O 103 - When Evaluation Meets the Need to Build up Resilient Communities: A Case Study on Women in Fisheries in the EU**

*A.C. Marangoni*<sup>1</sup>

<sup>1</sup> Coffey International, Evaluation & Research, London, United Kingdom

In March 2017, the European Commission and Member States of the Mediterranean and Black Sea coasts signed the MedFish4Ever Ministerial Declaration that established the need to “improve working conditions and promote the role of women in the fisheries sector and aquaculture”. While fishing and fish processing are male-dominated activities in Europe, women also play a key role in building up the resilience of fisheries communities by playing an invisible role in existing activities and the diversification of activities.

In the context of growing economic difficulties at the micro- and macro level for the fisheries communities, the contribution to resilience is a key dimension to policy-making. Although the contribution of women to the sector has gained visibility due to European policies on equal opportunities for women and men, there is only a very weak basis of evidence on women's actual contribution to the sector to inform policy-making.

The twin objective of gender equality and increased resilience of fisheries communities has led the European Commission to request a study on the presence and role of women in the fisheries sector. The study team faces a challenging environment where there is hardly any evidence available and where any data collection methods faces socio-cultural obstacles (resistance against the role of women in fisheries and against the recognition of the role of women in fisheries).

This case study will present and discuss the methodological choices made to collect evidence on the role of women and their presence in the fisheries sector, and highlight how they contribute to capacity building for more resilient policy-making. We will conclude that the neutrality of the study contributes to gathering relevant and good quality data, which in turn inform and support the formulation of policies to strengthen the resilience of fisheries communities.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 034 Integrating Gender Perspective in Evaluation

## O 104 - Impact Design for Women Empowerment and its Results

*T. Goetghebuer<sup>1</sup>*

<sup>1</sup> ADE, impact evaluation expert, Louvain-la-Neuve, Belgium

**Rationale:** ADE is conducting a rigorous impact evaluation of the Min Ajliki program in Morocco (<http://www.minajliki.ma>), promoting women entrepreneurship. Women are key actors to increase family welfare and develop their region, however, women entrepreneurship is constrained by limiting believes which can partly be overcome by supporting women empowerment. We succeeded to build a quasi-experimental design using a participatory approach involving 8 partners coming from public and private sector as well as from the civil society. Furthermore, data collection tools developed are tailored such as respecting ethics and local context, to enable to collect relevant and reliable indicators on women empowerment and welfare.

**Objectives sought:** This session aims to share ADE's main lessons learnt regarding participatory quasi-experimental evaluation design for a complex intervention in the women entrepreneurship field. It also aims at presenting the innovative data collection tool and the main results of baseline analysis of 1500 women survey to show where limiting believes constrain women entrepreneurship and how various types of partners team up to overcome these constraints.

**Brief narrative and justification:** The aim of the session is to promote learning and arise interest in the use of rigorous impact evaluation methodologies which are feasible on the field and relevant for the key stakeholders. The intervention is dealing with four different types of women beneficiaries in urban and rural area, we will describe how the design has evolved for each category due to partners and context constraints but remains rigorous (cluster randomization and phase-in design) and fits the real life of an intervention implementation.

Baseline data are being collected in April 2018 using an innovative process (e.g. group interviews producing individual data). The presentation will also briefly present this tool and the reason why it was the most meaningful tool for the women themselves as well as to meet the purpose of the study.

Finally, we will present the main findings about limiting believes hindering women entrepreneurship development in the Fès-Meknes Region in Morocco.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 033 Turbulent Contexts and Evaluation

## O 105 - Evaluation Governance in a Complex System: A Case Study of CGIAR Evaluations

*R. Sauvinet-Bedouin<sup>1</sup>, R. Birner<sup>2</sup>*

<sup>1</sup> CGIAR, Independent Evaluation Arrangement IEA, Rome, Italy

<sup>2</sup> University of Hohenheim, Social and Institutional Change in Agricultural Development, Hohenheim, Germany

The fifteen international agricultural research centers that form the CGIAR play an important role in reducing hunger and poverty by promoting agricultural innovations in the developing world. Together, the centers employ almost 10,000 scientists and staff and receive more than 800 million USD of funding from a consortium of diverse donors. The food price crisis of 2008 triggered a major reform of the CGIAR, which aimed at strengthening the development impact of CGIAR research. The reform led to the creation of thematic CGIAR Research Programs that cut across the fifteen centers and, hence, introduced a matrix management system in the CGIAR. The reform also led to the establishment of the Independent Evaluation Arrangement (IEA) of the CGIAR, which reflects an increased commitment of the reformed CGIAR to use evaluations as the basis for accountability, learning and continuous system improvement.

Since its establishment, the IEA had to operate under conditions of unpredictability and complexity. Unpredictability is inherent in the nature of research operations, but the CGIAR reform process added to the unpredictability, since it involved perpetuated changes in the governance and management of the CGIAR and a funding structure and cycle that do not match the requirements of effective research processes. The complexity of the institutional structure of the CGIAR was already high, but the system became even more complex as a consequence of the reform.

By analyzing the experience of the IEA, the proposed paper aims to contribute to a better understanding of the role that evaluation can play under conditions of complexity and unpredictability. The analysis focuses on the evaluations of the fifteen cross-cutting CGIAR Research Programs (CRPs), which cover almost the entire research portfolio of the CGIAR centers. IEA commissioned and carried out 10 CRP evaluations between 2013 and early 2016 and IEA supported the evaluations of the five remaining CRPs that were commissioned by the programmes themselves. IEA also published a "Synthesis Review of Lessons Learned of the 15 CRP Evaluations" in 2016.

The paper consists of three parts. In Part 1, the approach and methodologies that were applied to conduct the 15 evaluations and the synthesis review is described and analyzed in a comparative perspective with regard to their efficacy in addressing the different areas of evaluation, such as relevance and priority setting of the research programs, quality of science, impact and value added of the CRPs. The second part of the paper examines how the findings and recommendations of the CRP evaluations were communicated within the complex management structure of the CGIAR. Special attention is paid on how the evaluations were received by the management of the CRPs and the CGIAR Centers and by the funders of the CGIAR. The extent to which the recommendations were adopted and followed up is reviewed, as well. In Part 3, the authors analyze how the CRP evaluations contributed to the reform process of the CGIAR. Lessons are derived with a special focus on issues to be considered for evaluations that take place under conditions of unpredictability and complexity.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

#### S 044 Appreciating Capabilities

### O 107 - Using Theory of Change to Structure the Evaluation of the “Dynamics” of the Flemish “Focus on Talent” Policy

*E. Desmedt*<sup>1</sup>

<sup>1</sup> IDEA Consult, Regional development and labour market team, Brussels, Belgium

Since 2016, 'Focus on talent' is the new approach of the Flemish Government for a higher employment rate and labour market inclusiveness. A target group policy with specific measures is replaced by a policy with general measures, focusing on the strength of an individual. Also the role of the government changed: a bottom-up approach is applied, in which various actors receive project funding to further develop and implement the policy vision, with the aim of initiating a change in mentality in the labor market.

The purpose of this evaluation, requested by the government, is:

- Acquiring insight in how the policy works in the field, whether and to what extent the new policy creates the expected 'new dynamics'.
- Acquiring insight in the change processes themselves, to gain insight in the underlying mechanisms that contribute to the success of projects.

It is a mid-term evaluation, emphasizing the learning aspect.

The main challenge in the design of this evaluation is that the 2 central concepts in the objectives of the evaluation – 'dynamics' and 'mechanisms' – are little tangible and difficult to operationalize. We chose a theory-based evaluation approach to tackle this: drawing the theory of change opens the 'black box' of the evaluated policy and projects, and provides insight in the dynamics and the underlying mechanisms. It is also an approach in which the learning aspect is integrated: the drafting and evaluation of the theory of change must take place together with the stakeholders.

The government requested a case study approach, with a minimum of 4 cases. For each case, three steps are taken:

1. Drawing the theory of change. During a workshop, we investigate how the stakeholders hypothesize that the project will be working (from input to reaching the final impact) and what they think the underlying mechanisms are.
2. Evaluation. We test the theory of change against what can be empirically observed, on the basis of monitoring data, interviews and focus groups. Qualitative and quantitative methods will be flexibly used.
3. Consolidation. We organize a final workshop to present our draft case study report: can the stakeholders agree with the evaluation findings and conclusions? Preferably this is done in a mixed group: people responsible for drawing up the project, people responsible for its implementation, and also a representation of the target group.

In our contribution, we will present work in progress: at EES in October, 1 or 2 case studies will be completed, and the others will be in the start-up phase. We'll discuss the challenges we met in the first case studies (e.g. reactions of the stakeholders, issues with drawing the theory of change, availability of data, ...), how we try to take them on in the next, and what lessons can be learned about evaluation in general.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 036 Ethical Issues and Evaluation**

## **O 108 - Plagiarism Undermines Achievement of Quality Education in African Universities**

*O. John Michael Maxel<sup>1</sup>*

<sup>1</sup> Uganda Evaluation Association, member, Kampala, Uganda

**Introduction:** Monitoring and evaluation involve systematic checking to see whether the product or service meets the set standards. Globally Universities focus on system improvement, process, procedures and mainstreaming actions intended for the achievement, maintenance, monitoring and achievement of quality in the University. Monitoring and evaluation in the different entities have to systematically reflect the realities in the organizations. The different stakeholders in monitoring and evaluation have to exercise honesty for credible results.

However, plagiarism kept undermining these noble efforts. Plagiarism is evident in academic evaluation process reflected by the monitoring and evaluation professionals duplicating the previous evaluation reports. This can be seen in terms of the similar methodologies, approaches and the reports of the external evaluation duplicating the internal evaluation reports. This undermines the process of transformation of Africa as a continent by the production of half-baked evaluation results. This cannot significantly transform practice in of monitoring and evaluation. The reports cannot systematically improve the practice in the government and non-government organizations. The author undertook a qualitative and quantitative study to investigate the concept of plagiarism in University education negative vice slowly destroying fabrics of university education.

The paper argues that in order to build proper standards monitoring and evaluation associations have to incorporate honesty and integrity. This is by the development of appropriate mechanisms for elimination of plagiarism and all forms of dishonesty. The paper discusses the concept of plagiarism, causes of plagiarism and recommends strategies dealing with plagiarism and academic dishonesty. This is critical in holistically in an improvement of monitoring and evaluation practice.

**Justification:** Monitoring and evaluation are critical to understanding design, implementation and practice. This forms a basis for improvement in the practice as the process used reveals strengths and weakness of the design, processes, systems, and implementation.

However, plagiarism kept undermining these noble efforts. Plagiarism is evident in academic evaluation process reflected by the monitoring and evaluation professionals duplicating the previous evaluation reports. This can be seen in terms of the similar methodologies, approaches and the reports of the external evaluation duplicating the internal evaluation reports. This undermines the process of transformation of Africa as a continent by the production of half-baked evaluation results. This cannot significantly transform practice in of monitoring and evaluation.

Monitoring and evaluation association have undertaken several efforts to improve the practice of monitoring and evaluation but little attention has been given to plagiarism and dishonesty. The paper in the transformation of evaluation practice as it highlights the different forms of plagiarism and dishonesty. This paper is critical in supporting the professionals to adhere to the highest standards in monitoring and evaluation.

**Purpose:** The purpose of this study is to provide the impact of plagiarism in monitoring and evaluation practice by the different associations.

**Conclusion:** Plagiarism silently and systematically plagues quality standards of monitoring and evaluation practice by destroying innovation and creativity. In order to improve the quality of monitoring and evaluation practice; different associations have to incorporate honesty and integrity. This is critical in holistically quality assurance in Universities in Africa.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 013 Breaking the Boundaries of Evaluation

## O 109 - Towards a General Theory of M&E: An Approach when Evaluating Potentially Hostile and Dishonest Clients

*E.E. Goetsch*<sup>1</sup>

<sup>1</sup> University of the Witwatersrand, School of Law, Johannesburg, South Africa

**Rationale:** One expects “tell the truth, the whole truth and nothing but the truth” to be the standard for reports. In practice, few evaluators can afford to be whistle-blowers. Many prefer to play safe by avoiding awkward topics. This temptation is maximised in the international development aid sector. It is prone to financial frauds and false claims, leaving the evaluator with the dilemma of bridging the gap between over-promises and under-delivery. In a time of austerity, with tough competition for tight budgets, these gaps widen.

**Objectives sought:** To present an approach to evaluation and a model of reporting that can credibly serve the practitioner as a standard that protects his or her independence, integrity and financial security when the client who commissions the report also has something to hide.

**Brief narrative:** The approach involves a number of changes. Firstly, it expands the scope of the evaluation. It puts the organisation and its project cycle wholly inside the report: the money of the inputs, the staffing of the activities, the delivery of the outputs and the impact of the outcomes. It puts all 10 areas of management and their decisions and results explicitly inside the report as well. It does the same with the 3 tiers of the organisation and the 9 steps of the international development funding flow. This leaves nowhere to hide. Secondly, it reverses the onus of proof. Rather than the evaluator having to prove breach, the client needs to prove compliance. The model report covers the universe of issues, lists the areas and types of non-compliance and allows the evaluator to safely report whether the evidence exists that misconduct did not happen. It comes with examples of how agencies routinely “game the system” to cover up false claims.

**Justification:** The approach and the model offer a quality standard for a final report. It has many benefits. It helps evaluations help people to improve their lives and make our societies more resilient. It reduces unpredictability and complexity. It helps when designing and managing evaluations. It informs evaluation systems. It helps to rethink evaluation methods, design and criteria. It combines methods in evaluation. It addresses delivery risk in uncertain futures. It facilitates collecting and analysing data and reporting issues particularly in challenging contexts. It provides a dashboard that integrates ICT, M&E and managers. It enables evaluation to become foresight. It rewards flexibility and handles complexity. It develops the field of Evaluation to promote resilience and action in this critical time. It solves a challenge and offers opportunities for the evaluation field. It addresses some of the dilemmas and trends in professionalism, standards and ethical norms. It advances the theory and applies ethical values to evaluation. It promotes Evaluation Associations as custodians of a professional standard. It protects the independence of evaluators and our relevance and responsiveness. It supports the partnerships and stakeholders who make up the international development aid industry and greatly strengthens the communicating, using and embedding of evaluation. It is truly intersectional, since it connects M&E to every role-player in the organisation.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 036 Ethical Issues and Evaluation

## O 110 - A study on the Evaluation of Ethical Environment for Science and technology

*C. Fan*<sup>1</sup>

<sup>1</sup> Institutes of Science and Development- Chinese, Institute of Science Strategy, Beijing, China

Scientific research and technological development, such as gene editing, nanotechnologies, may bring about some ethical issues and related social and law issues. It is necessary to have a suitable ethical environment for the development of science and technology to meet the ethical requirements.

This paper 'aim is to study what constitute the ethical environment for science and technology development and how to assess it. It defines the ethical environment for the development of science and technology as the sum of various factors for making ethical value judgment and regulating the development of science and technology ethically, which is embodied in cognitive (social ideology, the mainstream values), institution (ethical rules, review mechanism, laws and regulations), organization (relevant organization rules and culture), society (the knowledge and interests of the social actors). The ethical environment of scientific and technological development includes the whole chain from scientific decision-making, scientific research, technological development, application and communication. The ethical environment is helpful for understanding, communication, solving and education of ethical issues of science and technology development. Thus, ethical environment plays a regulation role (inhibition, limitation, encouragement) in the development of science and technology.

The study identifies 11 factors as the main elements of the ethical environment: (1) the international scientific ethical norms;(2) mainstream values related to the development of science and technology;(3) national science and technology policy documents; (4) science and technology policy advisory mechanism ;(5) relevant laws and regulations;(6) ethical guiding principles and ethical review mechanism;(7) scientists' ethical awareness and social responsibility;(8) public awareness and participation in scientific ethics;(9) research on ethical, legal and social issues of science and technology;(10) media reports on the ethical issues of science and technology;(11) scientific ethics education.

Based on the above study, the paper proposes a framework for assessing the ethical environment, which uses the 11 factors as major indexes for assessing of ethical environment and set up assessment indicators. Limited by words, just give an example, for the Index (5) ethical guiding principles and ethical review mechanism, the assessment indicators are: are ethical guidelines established in important areas of science and technology, refer to international ethical standards? How effective is the implementation of ethical guidelines? Has an ethical review committee and mechanism established ? How does the ethics committee work? and did it work ?

This framework has two main USES: (1) evaluate the utility or function of the ethical environment (eg, will the ethical environment influence development of science and technology to meet the ethical requirements); (2) diagnose weaknesses and problems of the environment and make suggestion for improvement. (eg, did the ethical review committee work? how to improve it?)

By using the framework, we made an evaluation study on ethical environment of Nanotechnology in China comparing with U.S. It can be concluded that the ethical environment is enabling nanotechnology development ethically. But there is still some problems need to be solved. The framework could be used to evaluate for ethical environment of other emerging technologies

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 037 Planning and Budgeting for Social Change

## O 112 - Use of Technology to Improve Planning, Budgeting and Reporting: Lessons from Uganda's Program-Budgeting System (PBS)

*R. Obuku<sup>1</sup>, W. Josephine<sup>2</sup>*

<sup>1</sup> Amolatar District Local Government, Planning, Amolatar, Uganda

<sup>2</sup> Parliament of Uganda, Monitoring and Evaluation, Kampala, Uganda

Many African countries are following the worldwide trend of introducing a performance orientation into annual budget processes and of planning budget expenditures over a multiyear period. A major aim is to more closely align strategic socio-economic planning with annual budgets and medium term budgetary frameworks (MTBFs). In francophone Africa, program budgeting has been chosen as the preferred 'model' of performance budgeting while in other parts of Africa, although program-based budgeting is being introduced in some countries and there is a greater diversity in the emerging performance-based budget systems. The main objective of PBB is to link annual budget spending allocations to budget performance. During budget preparation, program performance is taken into account using systematic performance information.

Over 80 percent of African countries are introducing or are committed to introducing some form of Program based budgeting and, in many cases, Program based budgeting reforms are being adopted as part of a broader package of public financial management (PFM) reforms, pressure from regional bodies in Africa and the donor community. Two African countries (Mauritius and South Africa) have a functioning Program based budgeting system, and eight others (Ethiopia, Kenya, Liberia, Malawi, Mozambique, Namibia, Tanzania and Uganda) have made considerable progress towards Program based budgeting implementation

The Government of Uganda introduced several reforms to improve overall management of public resources which are in line with the existing legal framework (Constitution- 1995, PFMA-2015, Local Government Act) and these reforms were informed by studies done within and outside government, review of past performance and monitoring reports, local and international based practice while some are aimed at mitigating the existing challenges. In FY 2017/18, Government of Uganda transited from Output Oriented Budgeting to Performance Based Budgeting (PBB) which is a budgeting structure where Funds are distributed by program or functional area and based on the nature of the activities performed by the program A program brings together different interventions that have a shared objective often referred to as outcomes

In order to implement Programme Based Budgeting, in the Financial year 2016/17 the government of Uganda started to define the structures of Programme Based Budgeting (PBB), Logical Framework for both Local and Central Government as well as initialising designs of the online Programme Based System (PBS) to operationalise and effectively implement Programme Based Budgeting. The PBS was ready for use by the Central Government to prepare Budget Estimates for FY 2017/18. In addition, more effort was put on the design of the Local Governments program budgeting system of which to date it has become the official planning, budgeting, Monitoring, evaluation and reporting for government of Uganda projects and program

This paper shares this success story of Uganda's program budgeting system in implementing program-based budgeting for planning, budgeting, Monitoring, evaluation and reporting for government of Uganda projects and program.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 037 Planning and Budgeting for Social Change

## O 113 - Evaluation of Gender Budgeting Enforcement Measures: The Ugandan Experience

*J. Watera*<sup>1</sup>

<sup>1</sup> Parliament of Uganda, Corporate Planning and Strategy, Kampala, Uganda

**Rationale:** The global dream as provided for in 2030 agenda or the 17 Sustainable Development Goal (SDGs) demands leaving no one behind. One hallmark step towards achieving this aspiration specifically SDG 5 is gender and equity planning and budgeting. This is an approach of allocating and utilizing country resources while taking into consideration the different needs, interests and constraints of various categories of people without any discrimination and addressing any imbalances that exist.

Originally, gender budgeting was developed in Australia in the 1980s. The 1984 Australia's landmark gender budgeting initiative required government ministries and departments to analyse the impact of the annual budget on women and girls, with a focus on public expenditures. A number of other international agreements or declarations since then reaffirm this commitment to use government budgets for gender equality and women's advancement objectives.

**Objectives:** Gender and equity budgeting is now widespread, more than 80 countries have tried some variant of gender budgeting, although these initiatives vary from place to place and country to country.

This paper presents the Uganda's journey towards gender budgeting with an evaluation of what has worked and/or not worked based on the existing enforcement measures. It highlights the rationale for gender and equity planning and budgeting and shares the 2018 assessment report (certificate of compliance) focusing on lessons registered from enforcement of gender budgeting in Uganda.

**Narrative and justification:** Uganda is among the countries in the world implementing gender and equity budgeting. In the period between 1962–2003, planning and budgeting in Uganda underwent several reforms with a view to achieving budgetary efficiency, transparency, accountability and value for money in public expenditure management. While significant achievements were registered with various reforms and enacted laws, there was little progress made in ensuring compliance with issues of gender equity in plans and budgets. This gap triggered the need to incorporate Gender and Equity Certificate into legislation as an enforcement measure.

Consequently, the Public Finance Management Act (2015) requires the minister responsible for Finance to issue a certificate, certifying that the budget framework paper is gender and equity responsive and specifying measures taken to equalize opportunities for women, men, persons with disabilities and other marginalized groups.

This is also in line with the ultimate goal of the second National Development Plan of Uganda, which is to attain a middle-income status by 2020 through strengthening the country's competitiveness for sustainable wealth creation, employment and inclusive growth. Gender and equity planning and budgeting is therefore very instrumental realizing a just and fair Uganda wherein all Ugandans have equal opportunity to participate, contribute and benefit from national development efforts.

The 2018 assessment report of compliance to sector budget framework papers with gender and equity requirement for financial year 2018/2019 is the fourth in series and main focus of this presentation of how it has worked. The author hopes to draw critical lessons for conference audience.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 037 Planning and Budgeting for Social Change

## O 114 - Evaluation of Blending as an Aid Modality

*V. Coppens*<sup>1</sup>

<sup>1</sup> ADE s.a., Head of Evaluation Department, Louvain la Neuve, Belgium

**Rationale:** A key aspect of building resilient societies is promoting sustainable and inclusive economic growth. The SDGs and the Agenda 2030 recognize in this respect the importance of multi stakeholders' partnerships and of mobilising and sharing financial resources. Blending (i.e. combination of grants with loans or equity from public and private financiers) and investment facilities (IFs) managed by multilateral (e.g. International Financial Institutions) or bilateral development actors (e.g. European Development Financial Institutions) have hence increasingly become key modalities for development partners. Blending will for instance represent 10 to 15% of EU total aid in coming years. In the last ten years, EU grants for blending (€3,4bn) helped to unlock an estimated volume of €57,3bn of investments in EU partner countries.

**Objectives:** This presentation aims to explain the challenges of evaluating blending and investment facilities as an aid modality for development partners, and to present the methodological frameworks developed by ADE which proved useful to this purpose.

**Narratives and justification:** Evaluating blending or investment facilities as an aid modality entails specific challenges. Indeed, it requires analysis beyond assessment of performance of operations, on, for instance, the relevance and comparative advantages of this aid modality for a development partner.

ADE has identified the following key lessons from an evaluation point of view:

- A two-dimensional intervention logic addressing the objectives of (i) the aid modality and (ii) the underlying operations;
- A set of evaluation questions covering the two dimensions;
- An extensive inventory and typology of operations;
- Multiple sources to face the often limited availability of information on development results, including but not limited to longer-term operations;
- A consultative approach, including an active role for policy-makers and the blending mechanism / investment facility management.

These lessons are based on ADE's track record of evaluations on this subject, including:

- the Evaluation of EU Blending (incl. 7 regional investment facilities);
- the Mid-Term and the on-going End Term Review of the ACP IF managed by the EIB,
- the on-going Evaluation of the Facility For Euro-Mediterranean Investment and Partnership (FEMIP) managed by the EIB;
- the Evaluation of the Infrastructure Development Fund managed by FMO.

Lessons further build on ADE's other "Aid Modality Evaluations" it conducted, notably those of the channelling of aid by the European Commission through the UN family, and through the World Bank and EIB, the evaluation of the use of different transfer modalities in ECHO Humanitarian Aid actions, and several multi-donor Budget Support evaluations.

All these evaluations have been highly appreciated by policy makers and bodies concerned.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 037 Planning and Budgeting for Social Change

## O 115 - Evaluation in the Governance of and Through Social Investments

*K. Nordesjö<sup>1</sup>, D. Mukhtar-Landgren<sup>2</sup>, M. Fred<sup>3</sup>*

<sup>1</sup> Malmö University, Centre for Work Life and Evaluation Studies, Malmö, Sweden

<sup>2</sup> Lund University, Department of Political Science, Lund, Sweden

<sup>3</sup> Malmö University, Department of Global Political Studies, Malmö, Sweden

It has been argued that public sector organizations have shortcomings in terms of service design, delivery and accountability, and therefore are in need of new and innovative organizational solutions to handle deep-rooted social problems. In tandem with these problem formulations, demands on increasing the utilization of the skills of local welfare professionals are being raised, while still expected to maintain a high quality of monitoring, evaluation and control. In addition, there is a strong belief in using different forms of temporary institutional experiments, social innovations, pilots or projects to “try out” new solutions (potential *best practices*) which – if they are deemed successful – can be “scaled up” and “travel” to new contexts.

*The social investment perspective*, as advocated by organizations such as the OECD, the EU and the World Bank, is an example of the ambition to reconcile the somewhat contradictory ideas of trust and collaboration on the one hand, and control and value for money on the other. At the core of this perspective lies the notion that investing in social services, youth policies and health care is an investment, which will lead to reduced governmental costs in the future. Social investments are characterized by being (1) *temporary initiatives*, (2) *clearly demarcated* from every day public administration and (3) *a collaborative effort*. Moreover, (4) “efforts must be *evidence-based* or involve some form of *method development*”, and possible to (5) *monitor and assess* from a societal perspective as well as financial perspective. Overall, these social investment characteristics create tensions for evaluation practice.

#### 1. The tension between flexibility and control

The social investment perspective advocates flexible local government organizations allowing quick mobilization, collaboration across organizational borders, flexible resource allocation and the trust in social innovations to emerge bottom-up. Still, the importance of control and measurement is emphasized, both from an economic and a social perspective. This implies that evaluation have to support social investment initiatives through both formative and summative strategies, using both emergent and predetermined criteria.

#### 2. The tension between outcome/process measurements

Criteria of success for human service interventions are typically hard to predetermine and difficult to measure quantitatively, and there is seldom one best solution to a particular social problem. Still, a dominant evaluation approach of the social investment perspective is cost benefit analysis based on aggregated impact evaluations, in general a positivistic evaluation logic.

#### 3. The tension between local knowledge and standardized solutions.

Despite the long-term perspective of the general social investment perspective it generally organizes activities in short-term projects. In an evaluation use perspective, there is a tension working with project participants in a participatory manner towards single and double loop learning, and treating the project as a closed off experimental unit to produce instrumental evidence based knowledge.

Starting from these tensions, the aim of this paper is to gain a deeper understanding how evaluation is related to the governance of, and through, social investment. Empirical data are from an ongoing qualitative institutional ethnography of several Swedish and Finnish local social investment initiatives 2018–2021.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 038 Practical Approaches to Measuring Research Impact****O 116 - Handling Complexity to Assess Agricultural Research Impact in Developing Countries: Taking Stock of Lessons from the Impress Ex-Post Method***S. Mathe<sup>1</sup>, G. Blundo-Canto<sup>2</sup>, G. Faure<sup>2</sup>, L. Temple<sup>2</sup>, A. Toillier<sup>3</sup>, B. Triomphe<sup>4</sup>, E. Hainzelin<sup>5</sup>*<sup>1</sup> CIRAD, UMR Innovation, Yaoundé, Cameroon<sup>2</sup> CIRAD, UMR Innovation, Montpellier, France<sup>3</sup> CIRAD, UMR Innovation, Ouagadougou, Burkina Faso<sup>4</sup> CIRAD, UMR Innovation, Mexico, Mexico<sup>5</sup> CIRAD, Dg-dg, Montpellier, France

Over the last decade, societal pressure has increased on public research agencies to provide evidence that their research is contributing to the achievement of development goals. Agricultural research centres working in developing countries participate in innovation processes that are characterized by complexity and systemic challenges. In such settings, interactions with different actors, from policy makers, to farmers' associations, to technical services, to NGOs, to private companies and so on, are central to generate outcomes – practice and behaviour changes – and impacts. Outcomes can include changes in national and regional policies, scientific and technical collaboration between formerly non-organized actors, changes in agricultural management decisions, and so on. The impacts stemming from these interactions and outcomes are multiple and can go from improved incomes or food security, to increased employment opportunities, to ecosystem and human resilience, and, in some circumstances, to negative effects (e.g. increased reliance on pesticides, increased economic risks...). However, the innovation processes to which agricultural research centres contribute in developing countries encompass specific challenges linked to data availability, information systems, complex interactions among actors involved, and diverse economic, social and cultural factors and conditions. Evaluating such processes through standard ex post impact assessment methods relying on experimental or quasi-experimental approaches is neither quite feasible nor necessarily relevant. Rather, analysing and reconstructing interactions between actors and the outcomes of these interactions is of utmost importance to be able to identify, understand and measure impacts. However, few impact evaluations in agricultural research focus simultaneously on 1) process – mechanisms through which outcomes and impacts are generated and what was the contribution of research -, 2) interactions of actors in a system that generate these impacts, and 3) ultimate impact measurement. In order to contribute to methodological advancement and tackle this challenge, the French Agricultural Research Centre for International Development (CIRAD) developed a holistic, systemic and participatory ex post impact evaluation method called *ImpresS* (Impact of Research in Southern countries). Between 2015 and 2016, *ImpresS* was tested on 13 case studies representing a diversity of innovation processes to which agricultural research for development contributed. Focusing on an in-depth analysis of the impact pathways, the *ImpresS* method allows making cause-and-effect linkages explicit between its elements (inputs, outputs, outcomes and impacts) as well as measuring the impacts of the change process. This paper takes stock of the lessons learned from the application of the *ImpresS* method in terms of evaluating complex innovation processes in complex agricultural systems. The ability of the method to balance between navigating complexity, ensuring active participation of key concerned or impacted actors, and providing rigorous evidence is analysed, while responding to the challenges of evaluation in developing countries. A comparison with standard evaluation methods used in agricultural research is drawn and recommendations for methodological advancements presented.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 038 Practical Approaches to Measuring Research Impact

## O 117 - Evaluating Research Uptake: Methodological Lessons from the Evaluation of the UK Department for International Development's Humanitarian Innovation and Evidence Programme

*R. Lloyd*<sup>1</sup>

<sup>1</sup> *Itad, Organisational Effectiveness, Hove, United Kingdom*

The UK Government's Department for International Development (DFID) has developed an ambitious strategy and programme, the Humanitarian Innovation and Evidence Programme (HIEP), to support the transformation of the humanitarian system to better produce, use and apply evidence and innovation for the benefit of people affected by humanitarian crises, something highly relevant to Europe now as well as other parts of the world. Innovatively, an evaluation was commissioned to run alongside the programme for five years from 2013–18.

The HIEP programme represents a significant, likely unprecedented, mobilisation of resources for multi-disciplinary evidence and research in the humanitarian sector. It is a major test of whether and how donor-funded programmes can produce robust evidence aimed specifically at improving policy and in practice for the most operation-oriented, time-pressured interventions in some of the most complex, insecure and often data-poor contexts.

The evaluation has now completed the fourth and final phase and it is timely to share learning on combining methods which are becoming increasingly popular in evaluation generally and their relevance to the evaluation of research, evidence and innovation. The evaluation assesses the programme against a theory of change using a mixed-methods approach. It also addresses evaluation questions of relevance, value for money, effectiveness and impact.

Learning to date from the evaluation is relevant to the evaluation community for three main reasons:

1. Firstly, the subject being evaluated is of interest. The effectiveness and impact of research and evidence and in government's effectiveness in promoting its use and application in the humanitarian sector where research has traditionally been under-used is extremely complex.
2. Secondly, the methodology for the evaluation is an innovative combination of: assessment against a theory of change; contribution analysis; multiple case studies and; management assessment. It aims to provide an independent assessment of progress and results as well as to contribute to learning about ways of working to improve the current and future programme effectiveness.
3. Thirdly, the five-year time period of the evaluation is a model that is beginning to be used in other evaluations and this evaluation has valuable experience in the area. The experience of combining a role of providing an independent assessment as well as of providing support to learning within the programme will be shared. In addition, the process of undertaking a long-term evaluation process in the midst of both a humanitarian context which is by definition volatile and unpredictable as well as an organisational context of DFID which like all institutions undergoes change in that time period is one that is relevant to evaluators in and outside of Europe working in humanitarian crises and other contexts.

The evaluation team combines evaluation experience with expertise in humanitarian policy and practice, research, communication and organisational change. The evaluation's aim is to share its process and findings with the evaluation and humanitarian communities.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 038 Practical Approaches to Measuring Research Impact

## O 118 - Developing and Using a Rubric Model for the Evaluation of Prerequisites for Clinical Research in Sweden

*M. Bergstrom*<sup>1</sup>

<sup>1</sup> Swedish Research Council, Department of Research Policy, Stockholm, Sweden

**Background:** The Swedish Government has commissioned the Swedish Research Council (SRC) to evaluate the quality of clinical research conducted in the county councils that are subjected to an agreement between the Swedish Government and certain county councils (ALF-regions) concerning cooperation on medical education, clinical research and development of healthcare (the ALF agreement). The evaluation assignment for the SRC covers three aspects of clinical research which was evaluated separately and independently by international panels, in three parallel evaluations:

- Quality of the scientific production
- Clinical significance and social benefit of the research
- Precondition for clinical research

This paper concerns the evaluation of prerequisites for clinical research. The results from the three evaluations will be used in a performance based resource allocation model. (20 percent of the research budget is to be reallocated based on the quality performance in the different ALF-regions)

This is the first time that an evaluation is focused solely on the quality of the prerequisites for clinical research. Therefore it has been a challenge to develop an evaluation model that is transparent, equal and relevant for measuring prerequisites for clinical research in ALF-region that varies to a large degree in size and scope of research.

**Method:** The evaluation of the prerequisites for clinical research was based on four components:

1. Access to Research infrastructures
2. Time for research
3. Next Generation Researcher
4. Career models for clinical research

The components were assessed as to how they were implemented in the seven ALF-regions, in terms of creating good prerequisites for clinical research. In order to guide the assessment on how these components have been implemented and developed, an assessment matrix (rubric) was developed. The focus for the assessment is on how each ALF-region has developed the respective component of the organisation in the operations, and with regard to structures, processes and results. The rubrics has also provided the basis for the data collection, partly via self-assessment carried out by the managements of the seven ALF regions, and also via a questionnaire sent to research leaders in the respective ALF regions. Evaluation process

The evaluation was carried out by an international panel in three steps:

1. Performing a pre-evaluation of the prerequisites for clinical research in the ALF regions based on the regions' self-assessments and a survey directed to principal investigators in the ALF regions and by using the rubric
2. Local site visits and hearings in the seven ALF regions
3. Final discussions and assessments, where the panel categorized the ALF regions into one of the three categories according to the performance based resource allocation model.

**Lessons learned:** This evaluation approach has been very successful with regards to using a rubric model for assessing prerequisites in seemingly similar areas of clinical research within highly variable contexts, i.e. organizational settings. However, the process needs to be very transparent and guided all the way through to assure that the results can be seen as fair, valid and just by the stakeholders.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 039 Advancing Systems Thinking and Scale in Evaluation

## O 119 - How Might a Shift Towards 'Systemic' Evaluation Better Respond to Challenges of Complexity and Uncertainty in Turbulent Times?

*B. Schmidt-Abbey<sup>1</sup>, M. Reynolds<sup>1</sup>, R. Ison<sup>1</sup>*

<sup>1</sup> *Open University, Applied Systems Thinking in Practice ASTIP Group- School of Engineering & Innovation- Open University, Milton Keynes, United Kingdom*

### What is the focus of this paper?

This paper is aimed at exploring the application of systems thinking approaches to research and evaluation practice in organisations at the science/policy interface, and the challenges involved in introducing such approaches.

A key question is: How can more systemic evaluation practices be accomplished, giving justice to the complexities of the policies assessed being evaluated and multiple perspectives and stakes involved, to make evaluation more complexity-sensitive?

### Why is this of interest to the evaluation community?

The primary purpose of this paper is to contribute to collective reflexive practice concerning the epistemic and methodological challenges experienced, and how to overcome them. It seeks to expose the emerging theoretical foundation, the research question and hypothesis to early feedback from the professional community of evaluators.

Specifically, this paper seeks to:

1. to gauge possible leads of promising inquiry paths for the evolving methodology design for this research project, informed by evaluation practitioners and their experiences of the challenges involved in introducing systems approaches in evaluation practices;
2. to initiate an exploration of systemic constraints enacted by funders and commissioners to transformation of evaluation practices.

The motivation for this paper is to challenge the dominance of current mainstream evaluation approaches and methods employed in complex (social) policies and phenomena.

### How will the paper be approached?

The research will seek to test the notion of 'systemic desirability and cultural feasibility' as described by systems practitioner Peter Checkland of moving beyond a 'first-order science' research and evaluation tradition, to further advance opportunities for applying systems concepts and approaches in evaluation practices.

It will look at what obstacles are encountered and how could these be overcome. The starting hypothesis is that there are institutional barriers to the 'cultural feasibility' of introducing systems approaches to research and evaluation practices, as the mainstream way of conducting these continues to dominate and imposes important constraints to such opportunities.

Introducing systems approaches into evaluations exposes paradigmatic and epistemic tensions. The default way of doing research and evaluation sits uncomfortably with systemic approaches, which can implicitly challenge the conventional way of framing and approaching evaluation inquiries.

This experience resonates with Chris Argyris' and Donald Schön's 'theory of action', in which there is a gap between a 'theory in use' (the 'normal' ways things are done in organisations) compared to 'espoused theory' (what is claimed to be done in rhetoric – for example, evaluations of complex policies).

To move forward from this, a 'shift' will be needed. This calls for 'double-loop' learning to bring about transformational change in the organisational and academic culture in organisations, and to dissolve the contradictions and 'organisational defensive routines'.

How these shifts can be brought about in practice and how the obstacles can be overcome is subject of the emerging research project. The paper will present some initial hypotheses based on an initial scoping exercise and seeks to explore their salience and relevance collaboratively with the conference participants' own experiences from their respective practice.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

S 039 Advancing Systems Thinking and Scale in Evaluation

## O 120 - Phenomenal Evaluation Culture? Or Responding to Phenomenon-based experimental governance through more systemic and responsive evaluation

K. Lähteenmaki-Smith<sup>1</sup>, P. Virtanen<sup>1</sup>

<sup>1</sup> Finnish Innovation Fund, Sitra, Helsinki, Finland

The presentation proposed here builds on the experience of a debate process put forward by Sitra in 2017–2018, upon which a new, more phenomenon-based and systemic evaluation practice is put forward, as part of a proposal for governance renewal and innovation for the next Finnish government strategy.

Many of the trends influencing public sector leadership today are of direct relevance for evaluation culture and practice. Amongst these trends are for instance *decreased predictability and accelerating speed of change in public sector leadership; inter-connectedness of decision-making and the need for cross-sectoral collaboration; innovative leadership methods and practices* (including experimentation, social innovation and co-creation); *digital, more customer-oriented services, citizen participation and deliberation in service development; ecosystem thinking, between private and public sectors and civil society* (including new funding and investment methods, such as alliances, impact investments and social impact bonds); *transformation of representative democracy and territorial and social polarisation*.

The agenda for a more phenomenon-based governance model, which better responds to the trends referred to above, necessitates a fundamental re-thinking of evaluation methods, practices and cultures. Phenomenon-based policies approach strategic agenda-setting and implementation through *holistic, cross-sector thinking*, familiar from pedagogy all through the 2010s, and only recently having made a serious transition to policy-making, with themes such as *social inclusion, circular economy or inclusive and sustainable growth* becoming central themes on the government agenda. In our view these holistic and horizontal phenomena are also particularly well-suited for more mission-oriented policy design (e.g. Mazzucato 2014 and 2018).

This type of policy-design in our view brings about a need for evaluation framework, which is at the same time *experimental in nature* (at best RCT-based or research driven, e.g. basic income experiment in Finland), *strategically timely* (both ex ante and on-going, e.g. sustainability experiments), *multi-dimensional / perspective* (e.g. municipal employment experiments) and *learning-oriented* (e.g. small-scale circular-economy experiments).

Policy experimentation represents one of the key facets of this new systemic approach to governance innovation. The type of policy trials and strategic experiments explored are diverse, thereby necessitating a multi-method evaluation approach. Main types include open-ended, though result or even mission-driven **experimentation** (most in line with the aspirations of traditional thinking for evidence-informed policy), piloting for **early implementation** (e.g. inspiring and facilitating cultural change), piloting for **demonstration** (more testing for validation purposes than open-ended learning) and **operationalising** policy through experimentation (e.g. 'trailblazing').

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 039 Advancing Systems Thinking and Scale in Evaluation**

**O 121 - A Principles-Based Approach to Evaluating Results at Scale and Scaling Strategies**

*T. Wind<sup>1</sup>*

<sup>1</sup> *International Development Research Centre, Senior Program Specialist- Evaluation, Ottawa- Ontario, Canada*

In our turbulent times, when there are serious issues to address, we all wish to ensure interventions, policies, programs and innovations are getting to the right people and the right jurisdictions, with the right intensity of results. In the International Development Research Centre, we are concerned with preventing infectious diseases like Chagas in hotspot communities throughout Central America. Or expanding legal protection for survivors of gender-based violence in India. Or ensuring that commercially-viable, affordable food fortification is available to improve nutrition in rural communities throughout Tanzania. Because the research we fund addresses pressing human and environmental problems, it matters to us that evidence-based solutions get to the right scale of impact.

As such, we are turning to evaluating 'scale' as a characteristic of results, and strategies for getting to scale.

A typical starting place to learn about scaling strategies would be commercial and technology sectors, with numerous examples of technologies and innovations going viral, or seeing exponential growth. There is much to learn about scale and scaling from those experiences. However, when it comes to social, health, economic and political innovations, IDRC looked to experiences of organizations in the global South in scaling research to explore how we should think about scale, and ultimately, how to evaluate it.

This presentation will share a principles-based framework to designing and evaluating results at scale, built on those experiences.

The framework has four interrelated principles. While the word "scale" sometimes suggests a "bigger is better" mindset, the first principle points to criteria for determining what might be an optimal scale – where optimal isn't necessarily about getting as big as possible, but rather about being appropriate to need and context. The second principle has to do with who decides what the optimal scale is, and on what basis those decisions are made. The third principle is about the need to coordinate with other actors and influences, which stresses the need for systems thinking in design and evaluation. And finally, the last principle is dynamic evaluation – referring to the need gather and use appropriate evidence to support and assess scaling processes as they unfold.

This presentation will share IDRC learning about "Scaling Science", drawing out the relevance of these principles for evaluating the scale of results outside the global South. Because even though it might not be named explicitly, scale is probably an issue in almost any evaluation. It comes up under issues of relevance, effectiveness, reach, and sustainability.

The presentation addresses the second conference strand, particularly "rethinking evaluation methods, design and criteria" (particularly criteria) and argues that using a principles-based approach in evaluation can be an effective approach for dealing with flexibility and complexity.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 040 Rethinking Theories of Change in Evaluation

## O 122 - Combining Relevant Social Theories to Structure Effective Evaluation Practice: The Case of a School to University Transition Program in STEM

J. Owen<sup>1</sup>, P. Andrew<sup>2</sup>

<sup>1</sup> The University of Melbourne, Centre for Program Evaluation, Melbourne, Australia

<sup>2</sup> Program Review Consultants, Program Review Consultants, Melbourne, Australia

**Objective:** The purpose of this paper is to discuss the use of social theories that can be used to structure evaluations designed to deliver outcomes of significance.

Narrative. To focus the discussion the intervention we will consider is Uni Bridges (UB). Based in Australia, UB is an educational innovation involving a partnership between La Trobe University and 12 secondary schools, designed to improve the quality of STEM curricula (Science, Technology, Engineering and Mathematics) in these schools, and to encourage students to proceed to further education after their schooling.

The intervention can thus be described as complicated, as it is a multi-level multi-site program.

The **first social theory** we employed was the notion of sustained interactivity (SA), a relationship between evaluator and client which relies on the use of informal as well as formal mechanisms to transmit real-time information designed to assist decision-making about an intervention. We are concerned here with SA during the early development and consolidation stages of UB.

Sustained interactivity is implied in several evaluation approaches (models), for example Developmental Evaluation (Patton) and Rapid Response (Wholey). Despite the fact that these models are well developed, it is more difficult to find, *within them*, descriptions or examples related to the following issues: \*how do evaluators respond to the needs of decision-makers when their information needs are often immediate and subject to change as an innovation is rolled out?

\*which transmission mechanisms are effective in these circumstances? \*what kinds of information are of most use in these circumstances?

The **second social theory** we employed related to innovation adoption (IA); and the benefits of this theory in structuring actual feedback to program deliverers. One aspect of this theory applies to the characteristics of the innovation itself, and the success of otherwise that schools had in implementing each of these characteristics of UB during the early development/consolidation stages.

Thus, for the success of this study, the evaluators drew on two important aspects of social theory.

**Rationale:** It is our view that without recourse to these aspects the design of the study would not have delivered quality findings to influence the impact of the intervention. The use of relevant theories as set out in this case imply the need for evaluators to be made aware of a range of established conceptual frameworks through formal and informal training programs

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 040 Rethinking Theories of Change in Evaluation

## O 123 - Assumptions Based Evaluation

*O. Feinstein*<sup>1</sup>

<sup>1</sup> *Universidad Complutense de Madrid, Centro Superior de Estudios de Gestión, Madrid, Spain*

Assumptions Based Evaluation (ABE) is an approach that helps evaluators to explore systematically the key assumptions, implicit and explicit, in the evaluation of programmes and/or projects. It is a multi-disciplinary approach, combining insights and methods from different disciplines, dealing with incentives, risk and uncertainty. Theory Based Evaluation (TBE) is a particular case of the ABE approach which, in that sense, is more general than TBE. On the other hand, it is more specific than TBE as it includes a set of guiding questions that facilitates, like an algorithm, the application of the approach.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 040 Rethinking Theories of Change in Evaluation

## O 125 - Towards a General Theory of Monitoring and Evaluation: A Standard Theory of Change

*E.E. Goetsch*<sup>1</sup>

<sup>1</sup> University of the Witwatersrand, School of Law, Johannesburg, South Africa

**Rationale:** The quality of a project's Theory of Change is key to funding, managing, measuring and reporting it successfully. In practice, they are often imprecise, inaccurate, incomplete or inconsistent. This leaves managers uncertain, monitors confused, evaluators frustrated and donors dissatisfied. A standard would bring clarity to the profession, comfort to the donors, results to the beneficiaries and rigour to the discipline.

**Objective:** To propose a model Theory of Change that satisfies role-players and stakeholders as accurate, complete, consistent, equal and precise when describing any and every project in the international development aid sector.

**Brief Narrative:** The author has occupied or advised almost all the positions in the international development aid sector. Especially, he has advised the managers of the donor and implementation agencies in their day-to-day decision-making. The model presented satisfies the occupational requirements of 10 areas of management in all fields of development and across the funding chain. It is accurate, insofar as it captures the key performance indicators in their correct positions. It is complete, insofar as it stretches from the recognition of the problem to the report of the solution and the 9 steps from funding to delivery. It is consistent, insofar as it applies equally to each and every project, it is universal, since it applies in every development field. It is precise, insofar as it treats the quantitative and qualitative elements mathematically. It is equal, since it is independent of gender, race or person. It is inter-sectional, since it recognises the perspectives and requirements of all stakeholders and role-players in the international development aid sector. It is practical, insofar as it was field-tested in resource-poor environments.

**Justification:** The model expresses a quality standard for a universal Theory of Change. It has many benefits. It helps evaluations help people to improve their lives and make our societies more resilient. It reduces unpredictability and complexity. It helps when designing and managing evaluations. It informs evaluation systems. It helps to rethink evaluation methods, design and criteria. It combines methods in evaluation. It addresses delivery risk in uncertain futures. It facilitates collecting and analysing data and reporting issues particularly in challenging contexts. It provides a dashboard that integrates ICT, M&E and managers. It supports emerging technologies including social media, big data, location systems, cyber ethnography etc. It enables evaluation to become foresight. It rewards flexibility and handles complexity. It develops the field of Evaluation to promote resilience and action in critical time. It captures challenges and opportunities for the evaluation field. It addresses some of the dilemmas and trends in professionalism, standards and ethical norms. It advances the theory and applies ethical values to evaluation. It promotes Evaluation Associations as custodians of professional standards. It protects the independence of evaluators and our relevance and responsiveness. It highlights the partnerships and stakeholders who make up the international development aid industry and greatly strengthens the communicating, using and embedding of evaluation.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 041 Evaluation Systems**

**O 126 - Evaluation Systems and Procedures – Is Formalism the Flipside of Institutionalised Evaluation?**

*K. Attstrom<sup>1</sup>*

<sup>1</sup> Ramboll Management Consulting, European Policy, Saint Nazaire de Pezan, France

This paper aims to discuss and explore the procedural and formalised evaluation systems in place in many international institutions, by taking a critical perspective on the evaluation system and processes of the European Commission. It will use examples from a range of evaluations, to illustrate how and when the formalised evaluation system may be less conducive to high quality evaluations, which can have an actual added value in the policy cycle. It will look at the effects of the Better Regulation Guidelines in different steps of the evaluation process, from the development of terms of reference, to the evaluation process and dissemination, from the perspective of the evaluator.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 041 Evaluation Systems

## O 127 - Evolution of Croatia's National Evaluation System through Europeanisation

*M. Sumpor<sup>1</sup>, N. Starc<sup>1</sup>, I. Đokić<sup>1</sup>*

<sup>1</sup> *The Institute of Economics- Zagreb, Regional Development, Zagreb, Croatia*

The Republic of Croatia adopted the new Law on Strategic Planning and Development Management System in December 2017. In this way, evaluation became an official component of the governance system at the national level. Apparently, the Croatian Government and Parliament have set the arena for establishing a national evaluation system for all public policies across all levels of governance.

In this paper, the effects of the Europeanisation process on building Croatia's national evaluation system initiated in the early 2000s will be presented. The analysis consists of a document review and interviews of participating government officials and evaluation experts. Within the Croatian public policy context, the authors reflect on experiences in evaluation of national regional policy related strategic documents, programmes and projects. Based on the review of the evaluation process, effects of the Europeanisation at an early stage will be assessed. Evaluation is part of the policy and programming cycle in EU member states. In this regard, impacts on national policy formation and operational programming processes after gaining the European Union (EU) membership in 2013 are analysed. Croatia adopted the Evaluation Strategy in 2012 (update 2018), which is focused on setting up a national evaluation system for European Structural and Investment Funds (ESIF). The pros and cons of these processes in the past 15 years will be taken forward and analysed against the new legal framework.

The results of this research and corresponding recommendations can contribute to better functioning of the new evaluation system being set up in Croatia and provide guidance for participants in this process. Formal and informal contacts and interactions are growing in number, but better institutional linkages, formal educational opportunities and more intensive networking activities are needed to raise existing evaluation capacities. Also, the development of evaluation system in Croatia can provide useful insights for other countries that are either in the process of EU integration or in the process of formally setting up national evaluation systems.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 041 Evaluation Systems

## O 128 - The Role of Evaluation Partnership in Strengthening Mutual Accountability and Development Effectiveness

R. Hamza<sup>1</sup>, J. Pennarz<sup>2</sup>

<sup>1</sup> Ministry of Investment and International Cooperation, Evaluation Department, Cairo, Egypt

<sup>2</sup> International Fund for Agricultural Development, Independent Evaluation Office, Rome, Italy

Development cooperation experienced a paradigm shift during the early 2000's, from a relationship between donor entities and recipient developing countries, into a more integrated partnership. International agreements of effective development cooperation, from the Paris Declaration to Busan Agreement, recognized **evaluation as a significant tool for improving development relations and results**. There was increased attention to institutionalizing, strengthening, and utilizing national evaluation systems and capacities to improve the effectiveness of development activities. **This paper presents a case where an evaluation partnership between the Government of Egypt and the International Fund for Agricultural Development (IFAD) has strengthened mutual accountability and development effectiveness.**

In Egypt, the Government established a national Monitoring and Evaluation Department within the Ministry of Investment and International Cooperation (MIIC). The Ministry is the responsible entity for planning, coordinating and overseeing development partners' activities towards the achievement of the national and international development agendas, Egypt Vision 2030, and the Sustainable Development Agenda 2030. Within this framework, the Evaluation Department is tasked to develop and implement a government-led M&E system that delivers data and knowledge to inform decision making and to improve the effectiveness of development activities.

Throughout 2016 and 2017, MIIC and IFAD set a positive example in leveraging their partnership for effective development cooperation, where the principles of ownership, transparency and mutual accountability were accentuated. The Evaluation Department at the Ministry and the Independent Evaluation Office (IEO) of IFAD conducted two parallel evaluations prior to the renewal of IFAD's third Country Strategic in 2018. The purpose of this activity was to provide the necessary evidence to guide the design of the Strategy through taking stock of the accomplishments of previous programs; capitalizing on achievements; learning from pitfalls, addressing challenges and documenting lessons learnt.

**This paper presents the successful experience of the two institutions in managing, synchronizing, and complementing the evaluation activities of one another.** Both evaluations used a combination of quantitative and qualitative methods. However, the IEO placed more emphasis on the overall country strategy and integration of the country program. It also covered non-lending activities (policy engagement, knowledge management, partnerships). As a result, it highlighted the need for IFAD to capture knowledge from experiences and to build a wider range of partnerships.

The Evaluation Department at MIIC, on the other hand, conducted a thematic evaluation that focused on the economic and social effects of different interventions within each focus area including; rural finance, irrigation, rural marketing, research and development, and gender. It found that IFAD activities tended to consider women as homogenous group overlooking the different characteristics of the targeted women, especially with regards to education. It also noted the absence of situation analysis and needs assessment studies which compromised the effectiveness of targeting.

The findings and recommendations of both reports facilitated an informed policy dialogue among all involved stakeholders that produced concrete road map to guide future decisions, and to improve the effectiveness of IFAD activities.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 042 Taking Stock of Development Evaluation**

**O 129 - Drawing Lessons from the Afred Database and the Implications for the Supply and Demand of Evaluations in Africa**

*L. Smith*<sup>1</sup>

<sup>1</sup> CLEAR AA- Wits University, Director, Johannesburg, South Africa

The purpose of the submission is to make a contribution to the field of Evaluation by presenting a regional analysis of evaluation practice in Anglophone Africa. The contribution draws on the findings from a book compiled by CLEAR AA based on the analysis of the African Evaluation Database, which draws on 2635 evaluations from 12 countries over ten years (2005–2015). The presentation will summarize key findings from the three sections of the book focusing on evaluation practice, methods and use. The first part of the presentation will draw on key findings related to evaluation practice and will speak to the challenge of a systems approach to capacity development on the continent in the absence of a common framework to define what core competencies are needed for the field. Part of this difficulty is linked to findings from analysis of the database that donors continue to dominate the demand for evaluations and disproportionately continue to commission Northern based consulting firms and academic institutions to carry these out. The epistemological frameworks driving these evaluations are in large part divorced from the local contexts and cultural narratives of where these evaluations take place. The second part on methodology, will delve into methodological trends of evaluations in 12 countries over a decade. It will speak to the challenges associated with who drives quantitative methods on the continent, the preponderance of RCTs and the questionable usefulness of these rigorous methods for contributing to strategic reflections on program reforms. This part of the presentation will also draw on the limited application of quality frameworks by donors commissioning evaluations for standardizing qualitative approaches to evaluations. The third component will focus on the tensions between purpose and use based on a sampling of 121 evaluations from the database. The findings reveal the dominance of evaluations being used for the purpose of management, which fosters greater compliance at the expense of using evaluations for the purpose of governance. Conclusions on this sample reveal that even though evaluation in the region is growing, the purpose of evaluation is still very closely tied to monitoring and accountability and therefore losing the opportunity for learning.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

#### S 042 Taking Stock of Development Evaluation

### O 130 - Status of Development Evaluation in Ethiopia: A Desk Review

*F.H. Reda*<sup>1</sup>

<sup>1</sup> *Ethiopian Evaluation Association, Strategic planning committee, Addis Ababa, Ethiopia*

Development evaluation has become a familiar public management tool in the modern world we live in. It has brought development issues of varying nature into the forefront of policy making. The practice of development evaluation has a wider scope depending on institutional strength, professional capacity and evaluation policy. This research paper aims to assess the status of development evaluation in Ethiopia using a desk review. All available researches on the subject of study are analyzed to highlight the level of maturity with regard to public evaluation culture in the country. The paper adopted the approach of Furubo and Sandahl (2002). These authors have developed nine indicators in which they give a score using a rating scale of numbers from 0 up to 2. They did this to evaluate the level of maturity which must be fulfilled to determine establishment of public evaluation culture in a country.

The result of the analysis depicted Ethiopia has performed well with regard to existence of national discourse concerning evaluation and availability of professional associations. The country has fallen in the middle with regards to degree of institutionalization – government, pluralism of institutions and proportion of outcome evaluations in relation to output and process evaluations. The analysis revealed conduct of evaluation in many policy domains, supply of domestic evaluators in different disciplines, degree of institutionalization – parliament and use of evaluation in the supreme audit institution does not exist. This research recommends that the country needs to have a national evaluation policy in order to address gaps.

My research paper will enlighten participants with the progress made on National Evaluation Capacities commitments since 2015 in Ethiopia. I will use different assessment criteria to show case the practice of development evaluation in Ethiopia. In doing this, my research paper informs conference participants the level of evaluation culture in Ethiopia through reporting the progress made on strengthening enabling environments, development of national evaluation policy frameworks, establishment of institutional arrangements, and the practice and use of development evaluation in public sector governance and policy making process.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 042 Taking Stock of Development Evaluation

## O 131 - How is it Possible to Design and Conduct Developmental Development Evaluation?

*M. Rökköläinen*<sup>1</sup>

<sup>1</sup> *The Ministry for Foreign Affairs of Finland, Development Evaluation Unit, Helsinki, Finland*

This paper is intended to provoke discussion on **different purposes of developmental evaluation**.

Development evaluation is a generic term for evaluations conducted in developing countries, usually focused on the effectiveness of aid programs and initiatives. Evaluation on development policy and cooperation is seen as a key element of transparency, good governance and an important contribution to strengthening democratic processes.

Developmental evaluation approach is often conducted under conditions of complexity with the focus on adaptation to local context. Developmental purpose, utilization focus, co-creation and innovation niche are the essential principles of developmental evaluations.

The Ministry for Foreign Affairs of Finland (MFA) carries out comprehensive, policy level evaluations (centralized) and project evaluations (decentralized). The purpose of the centralized evaluation is to serve strategic planning and decision making as well as learning from results. Mostly evaluation itself is defined to be tool for accountability and ensuring the policy priorities or cross-cutting topics. Scope of the evaluations has often defined to be formative, enhancing participatory approach and promoting learning of MFA and partners.

MFA commissions meta-evaluations on decentralized evaluation reports (latest 2015 – 2017). The rationale behind meta-evaluations is that they are valuable tool for accountability and to improve transparency towards partner countries. The purpose is to synthesize the results and to enhance the quality of decentralized evaluations as well as to derive recommendations how to improve MFA's development cooperation.

MFA takes regularly part in OECD/DAC peer reviews. One of the results of the latest review (2017) was that Finland would need to demonstrate how it contributes to partner countries' own national development goals and makes full use of country results. The recommendation was to use results information for multiple purposes (accountability, communication, direction, and learning) at multiple levels (corporate, sectorial, project).

Finland's approach to working with the private sector at the field of development cooperation has brought developmental approach to evaluation. The idea of conducting developmental evaluation throughout the Business with Impact -program (BEAM 2015–2019) is to enhance collaboration among partners and encourage co-operating with the business sector. The evaluation provides means to verify achievements against intended results as well as unintended consequences.

There is a trade of between different purposes of evaluations. This raises a question how it is possible to use results obtained from one context and for one purpose for other purposes? Meta-evaluations raise a question how to combine local (project level) and national targets. There is also trade-off between validity and reliability in evaluations in different contexts. Quite often there is also a need for quality assurance. It is important to clarify what is the purpose and who is the audience for the results of the evaluations. Without clarity of purpose, the tendency of all evaluation systems is to collect more and more data yet without knowing how to use data for developmental purposes. All tough it is important to recognize the evaluation processes and structures that further partnering, utilisation of results and leads to innovations, one critical question is how can developmental evaluation serve accountability needs and demands?

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 043 Values and Rights Based Approaches**

**O 132 - Supporting Responsiveness and Reciprocity: Changing Our Lens from Needs-Based to Rights-Based Evaluation**

*L. Bremner<sup>1</sup>, L.E. Lee<sup>2</sup>*

<sup>1</sup> Proactive Information Services, Presidnet, Sandy Hook, Canada

<sup>2</sup> Proactive Information Services Inc., Vice-President, Sandy Hook, Canada

This paper argues for a change of lens to move from needs-based to rights-based evaluation. The shift is needed so evaluators can respond to national and global contexts by changing attitudes and approaches to become responsive and inclusive. Many evaluations focus on programs whose 'beneficiaries' (individuals, communities) are seen as having needs or deficits, rather than recognizing and building on their existing assets. Often the voices of the evaluation 'beneficiaries' or stakeholders are muffled by the attitudes of those who have power and privilege leading to the use of exclusionary, paternalistic and inauthentic practices which need to be recognized as unethical in evaluation practice. UN conventions the Rights and Dignity of Persons with Disabilities the Convention on the Rights of the Child and the Declaration on the Rights of Indigenous Peoples, all speak to the right of people have to full participation in society. The implication is that the process is reciprocal; expertise and knowledge are shared between evaluator and beneficiary. Both contribute to mutual learning and understanding. This rights-based approach is also embedded in the global work of EvalPartners, such as EvalIndigenous, Gender Equity+, and EvalYouth. As professionals we must to find authentic ways of giving voice to and sharing power with those most affected by the programs we evaluate, thus recognizing the rights of all to meaningful participation. As evaluators, we are compelled expand our future to one that is inclusive both in terms of voices and methods if we are to address the crucial social, environmental, and economic issues that we face in today's turbulent world.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 043 Values and Rights Based Approaches

## O 133 - Common Challenges on Complexity, Intersectionality and Inclusive Evaluation Practices (or Questioning the Myth of the Impartial Evaluation)

A. Azevedo<sup>1</sup>

<sup>1</sup> *Oxfam GB, Planning- monitoring- evaluation and learning team, Oxford, United Kingdom*

Political economy thinking is the new goal on international development, but concrete challenges persist on how to translate that into monitoring, evaluation and learning practices. With few exceptions, these functions are not embedded in development organisations' programmatic cycles – learning and evidence are still not an integral part of the thinking as we evolve to adaptive systems.

Evaluation design aims, more and more, to embrace complexity, but putting complexity principles into practice can be challenging in the context described above. To get evaluation stakeholders to embrace new approaches to evaluation is one of the barriers evaluators find on their day-to-day work. Meanwhile, complex-appropriate approaches and thinking should be promoted for the questions they bring to the evaluation practice: issues of inclusivity and intersectionality being two key aspects of it. These are concepts highly valued by rights-based organisations and can offer us a great entry point for discussions on complexity.

I argue that intersectionality and inclusiveness come hand-in-hand with complexity. Both offer us means to think evaluation in a way that embraces non-linearity and systems thinking by posing questions on how to do 'evaluation from the margins'. But how can we facilitate these processes from thinking to practice as evaluators and evaluation commissioners in a donor-driven environment where scarcity of resources and time push us to linear approaches that are not capable of representing the lived experience of the people we work with?

This communication aims to explore this topic through reflecting about the role of evaluators as knowledge brokers and activists in discussions about complexity and intersectionality. I argue, with examples of previous experiences as an evaluator and evaluation commissioner, that evaluation professionals should have an active, political role in pushing for intersectional and complex approaches on their work by calling organisations on their own rights-based frameworks and the global Agenda 2030 of "Leaving no one behind". This call to 'think and act politically' on evaluation requires acknowledging evaluation as a political process of knowledge production. With this I aim also to hear from other evaluation colleagues on their experiences on promoting intersectional and complex approaches and how we, as champions of these two agendas, can find 'quick wins' to push these agendas forward.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 043 Values and Rights Based Approaches

## O 134 - Bridging the Divide: The Relationship Between Gender-based Analysis “Plus” (GBA+) Use and its Quality

*J. Whynot<sup>1</sup>*

<sup>1</sup> University of Ottawa, Student, Ottawa, Canada

This discussion reframes Canada's federal government conceptualizations of gender-based analysis plus (GBA+) beyond its current definition as an analytical process (SWC, 2017) to include enhanced interpretations most importantly for this discussion as an output, a tool, and a competency (Whynot for ESDC, 2017). Given the nascence of full scale GBA+ implementation, the role of quality, and value of GBA+ inclusion across the results domain for inclusion across the policy cycle have yet to be systematically assessed. Preliminary findings from case study research are shared highlighting conceptual, instrumental and symbolic uses found in participating departments.

As a process, the application of an intersectional lens to federal government spending aligns itself to hearing various interpretations of truth. Sex/gender serve as entry point analytical variables, with additional identity dimensions (such as age, religion, language, income, sexual identity) reframing homogeneous perceptions of the beneficiaries served by government funded interventions. The federal government evaluation function serves as one GBA+ production mechanism, other functional producers of GBA+ include program performance measurement, policy and research units (UNWomen, 1995).

Canada has had a longstanding commitment to integrate GBA+ in both the evaluation function, and within the policy cycle (SWC, 2017); but only with recent political leadership has this commitment been systematically enacted upon (Whynot, 2017). The consistent and comprehensive application by all government departments and agencies now requires the inclusion of GBA+ information at key policy cycle junctures including budget exercises, Treasury Board Submissions, and Memorandum to Cabinet. Spending requests now require the inclusion of GBA+ specific information to demonstrate that truths held by diverse groups of individuals have been considered for their potential impacts and mitigation strategies explored. Rarely will the “not relevant” option be selected, and proceed unchallenged.

Beyond GBA+ policy cycle timing limitations, a dearth of related competencies, and a scarcity of expertise, one of the main challenges lies in the clear definition of what constitutes a quality GBA+. Reports from Canada's Auditor General (2009, 2015) assessed GBA quality based on three criteria that included: reviewed relevant data sources, obtained perspectives of gender groups, and examined all GBA considerations. Subsequently issued guidance documents employ descriptors such as “sound”, “well-reasoned”, and “complete” to describe GBA+ output quality. Individual departments across government have drafted unique tools based on procedural elements identified in guidance documents resulting in different interpretations of GBA+ quality. The lack of a common understanding of what constitutes a quality GBA+ doesn't facilitate its acceptance by decision-makers who rely on quality information despite political influence to integrate GBA+ across the policy cycle. This potential perverse effect could conceivably end up negatively serving historically viewed alternative truths.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

#### S 044 Appreciating Capabilities

### O 135 - The Impostor Phenomenon In Evaluation Practice: Scholarly Analysis And Practical Solutions

*J. Lavelle<sup>1</sup>, N. Jones<sup>2</sup>, S. Donaldson<sup>2</sup>*

<sup>1</sup> University of Minnesota-Twin Cities, Organizational Leadership- Policy- & Development: Evaluation Studies, Minneapolis, USA

<sup>2</sup> Claremont Graduate University, Social Science- Policy- & Evaluation, Claremont, USA

The impostor phenomenon (IP) refers a range of negative emotions associated with perceived fraudulent competence in a given field or with a specific set of skills that the person must use to be successful with his or her trade. The phenomenon is common, and scholarly research suggests that close to 70% of all people feel like impostors for some part of their career (Gravois, 2007). Indeed, many professionals can relate with the perception that they are not wholly qualified for their current position, that someone else must be imminently more qualified to handle their work tasks, or that they really need to improve their skills in an effort of self-protection and preservation.

In the field of evaluation, IP is an experience to which that many practitioners and scholars can relate. Indeed, this might go hand-in-hand with the fact that many evaluators “fall in to” evaluation, and so we might be left with a sense of non-belonging, or a sense that we are simply playing the role that was intended for someone else. Many students and professionals alike have wondered aloud “who are we to call ourselves evaluators, what do we really know about how to make judgements of merit, worth, and significance?” We call this emotional state EIP: evaluator impostor phenomenon.

Drawing from the scholarly literature in psychology, business, and counseling, we suggest that EIP is a topic that can be explored, discussed, and normalized. Broadly, the factors that contribute to EIP may be sorted into two broad categories. The first relates to individual personality traits. These are useful for individuals interested in recognizing their traits that might lead them to be susceptible to EIP, or for evaluator educators interested in identifying which of their students might need additional assistance working through emotions related with EIP. The second category involves contextual elements, which are helpful for identifying under which conditions someone might experience an EIP episode, and prepare them to take steps to interrupt the negative emotional cycle. A third area of interest are in the demographic characteristics that might moderate the interplay between traits and contexts.

This paper summarizes the literature on the impostor phenomenon, applies it to the field of evaluation (EIP), identifies individual and contextual factors, and proposes a set of reasonable interventions to help lessen EIPs detrimental effects in the context of evaluation practice.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

#### S 044 Appreciating Capabilities

### O 136 - Integrating the Positive Thinking Approach in Development Evaluation to Deal with Complexity and to Make Local Actors more Resilient

L. Fantini<sup>1</sup>

<sup>1</sup> Independent consultant, None, Rome, Italy

The paper explores approaches in evaluation that are able to embrace the complexity of themes, contexts and relationships that characterize the development initiatives in the Global South rural areas and that are able to generate explicative knowledge about the effects produced (or not) by an intervention. In the pluralism of evaluation approaches, the paper focuses the attention on the Positive Thinking (PT), a family of approaches mainly used in the organizational learning facilitation and evaluation processes.

Those approaches have in common to be asset-based and strength-based, to propose a learning process that investigates the specific context starting from the existing successful practices by assuming that in every community, organization, or group something works. Moreover, they all imply a constructive way to conduct and report evaluation that contributes to find, suggest and/or create spaces to build a future change. The PT could be considered a change of mindset, a change of perspective through which to look at the specific context and intervention.

This change of mindset aims to overcome several cognitive and operational biases as well as established reductionist models in place in designing, managing and evaluating development interventions. It aims to deal with complex and adaptive systems by making explicit and recognizing value to their existing 'positive practices', to the role of the local actors as well as their capacities to be resilient, to bring innovative solutions and to be agents of change.

To evaluate project and programmes in the development field *wearing the lens of the positive thinking* could mean to question the predetermined conceptual frameworks, the theories, the models and objectives that frame the development intervention itself and to shift the attention to the real change produced considering the values, the capacities and the resources available in the local context.

This is the case in particular in evaluating development interventions that aim to strengthen resilience of rural poor. Resilience is an adaptive and locally contextualized concept. Realities of rural poor are very complex, and influenced by several external conditions that make them extremely dynamic and unpredictable. Understanding those factors is often more challenging for development planners and evaluators than for rural people themselves. Local actors in rural areas have the capacities to analyse, to map and to represent their realities, to propose solutions, to forecast, to estimate, to compare, to score more than planners and evaluators expect.

The paper demonstrates through a case study how the PT approach can support to make explicit the existing processes and capacities, the emergent dynamics and the relationships among the actors part of the system, remained unexplored and/or underestimated.

The paper concludes arguing through specific examples of field work how an evaluation design based on PT approach is in line with the south-south and triangular cooperation approach as an alternative key mechanism of aid delivery to tackle the new SDGs making local actors full-fledged development partners.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 045 Evaluating Health Programs

## O 138 - Understanding Usage Complexity in Evaluating Stock Visibility Solution – an mHealth Intervention in South Africa

*J. Govender<sup>1</sup>, V. Skiti<sup>1</sup>, G. Morar<sup>1</sup>, N. Ryavudzo<sup>2</sup>, R. Awasi<sup>3</sup>*

<sup>1</sup> Data Innovator, Johannesburg, South Africa

<sup>2</sup> National Department of Health, Johannesburg, South Africa

<sup>3</sup> Vodacom, Johannesburg, South Africa

**Rational:** The use of mobile technologies to track and improve health outcomes, referred to as mHealth, has been increasingly adopted and is a rapidly expanding practice[1]. In 2014, catalyzed by civil society voicing concerns around ARV stock levels, the South African National Department of Health deployed such a mobile application. This application – stock visibility solution (SVS) – with the goal to improve the use of stock management data to take appropriate proactive action when needed in the tumultuous context of managing HIV and TB. In 2017, an evaluation of this national application was commissioned.

**Objectives sought:** The primary aims of the evaluation were to determine how and to what extent do various managers in the supply chain utilize SVS medicine availability data in their decision making and supply chain management processes; identify enabling factors for the optimal use of the SVS for reporting medicine availability at facilities; and explore end-user perceptions and attitudes. As evaluators, we aimed to take a unique approach by drawing from evaluation and software user experience theory.

**Narrative and justification:** The evaluation of mHealth solutions is complex, in that it requires new approaches to assessing not only the data usage outcomes, but also the user experience and technical factors which influence data quality and use. Guided by the programme theory and using the Quality in Use Metrics (QUIM) adapted from software user experience approach, an evaluation approach was adapted to use, to study, to define, and to validate quality in use of data generated from SVS. The specific factors assessed include: effectiveness, efficiency, satisfaction, productivity, safety, accessibility. In addition, we conducted interviews at every data use level as well as secondary analysis of the SVS data. 17 interviews were conducted, 14 surveys, and secondary data analysis.

The findings eluded to challenges and lessons at various levels, including policies, human capacity and skills, documentation and non-standardization of processes and technology specific issues. The integration of assessing user experience, as it identified clear bottle necks in how the application was used, and interpretation of design which directly affected quality of data and the use in planning and decision-making around essential drug management. For example, the changing of the loading icon from green to grey raised confusion among users, who felt less confident that their data had been uploaded. Users also indicated not being aware of how to generate reports on various districts, and instead logged in in many profiles. These seemingly minor concerns are critical hindering factors in settings where technology literacy is low. It prevents intervention success at the low levels of the programme theory, significantly reducing chances of the expected outcomes.

This evaluation presented critical findings to improve the national SVS management, and also presents an important approach to assessing usage of mHealth platforms as a short term outcome, in addition to longer term outcomes in data usage in decision-making.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 045 Evaluating Health Programs

## O 139 - Expediting Continuous Evaluation Models as Suitable Early Warning Systems for Measuring Impact of Health Programs

*J. Magoola Okalangh<sup>1</sup>, S. Wadiembe<sup>1</sup>, A. Kunihiro<sup>1</sup>, A.C. Barbara<sup>1</sup>, F. Nalubega<sup>1</sup>, P. Gorrell<sup>1</sup>*

<sup>1</sup> Social & Scientific Systems Inc., USAID/Strategic Information Technical Support, Kampala, Uganda

**Rationale:** Over the last decade, Uganda with support of key development partners in United States Agency for International Development (USAID), The Centers for Disease Control & Prevention (CDC), Global Fund, UNAIDS & United Nation's Population Fund (UNFPA) among others have worked diligently towards developing the country's systems for monitoring and evaluating health impact. These numerous development partners heavily committed and invested towards developing the "one M&E system" ministry of health concept as they fostered the birth and development of Uganda's Health Management Information System (HMIS) and electronic District Health Management Information System (DHIS2). Vital health impact indicators have been identified and are periodically tracked and measured through DHIS2, Uganda Demographic Health Survey (UDHS) Uganda Population-Based HIV Impact Assessment (UPHIA) and most recently through the Lot Quality Assurance Sampling (LQAS) methodology.

**Objectives:** To establish the level of effectiveness of the continuous evaluation methodologies as early warning techniques for measuring long-term health impact

### Brief narrative and justification

**Methods used:** This paper draws comparative analytical of key health impact indicators drawn across the different approaches and techniques used in estimating health impact. The analysis triangulated 2016 results on key health impact indicators from four major methodologies used in measuring health impact in Uganda. The four techniques compared are; the District Health Information System (DHIS2) which is Uganda's cornerstone of routine and periodic data management, the Uganda Population-Based HIV Impact Assessment (UPHIA), Uganda Demographic Health Survey (UDHS) and most recently the Lot Quality Assurance Sampling (LQAS) which utilizes the smallest batches of the target population to evaluate health impact periodically. With the exception of the LQAS which did not cover two of the 10 regions, the other three methodologies covered all the 10 regions of Uganda; Central 1, Central 2, Kampala, East-Central, Mid-East, North-East, West Nile, Mid-North, Mid-West and South-West. LQAS.

**Results:** In 2016, data from three methodologies; UPHIA, DHIS2 and LQAS revealed that the HIV Prevalence at community level has a directly proportional trend across all the 10 regions of Uganda. While the UDHS showed the HIV Counseling & Testing (HCT) Coverage at 52.5%, the LQAS registered this at 54.7% across all the regions of Uganda. The use of modern family planning methods among sexually active women aged 15–49 years indicated 37.4% and 34.6% in the LQAS and UDHS respectively. Delivery with professional care registered 70% in the LQAS and 73% for UDHS report.

**Conclusions:** The LQAS methodology yields results within acceptable scientific standards with the other three major methodologies used to track and measure health impact in Uganda. Therefore, this technique proves to be an effective early warning mechanism for measuring of health impact.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 045 Evaluating Health Programs**

**O 140 - The Evaluability of Small-scale Community-based Approaches to Obesity Prevention: Pragmatics, Purpose, Power and Politics**

*E. Gadsby<sup>1</sup>, S. Hotham<sup>1</sup>, R. Merritt<sup>1</sup>*

*<sup>1</sup> University of Kent, Centre for Health Services Studies, Canterbury, United Kingdom*

Reducing levels of obesity has become an urgent issue for public health in the UK. However, in the context of ever-tightening purse-strings, local authorities are having to make very difficult decisions about how to prioritise interventions, and are ever mindful of good evidence to demonstrate the value (or potential value) of their investments. Whilst there is an increasing understanding of the need to take a 'whole systems' approach to tackling obesity, there is currently a lack of clarity about what this means in practice, particularly at a local level. In an effort to bridge this gap, a local authority in England decided to design and pilot a small-scale community-based 'whole system' programme, and to commission an external evaluation of it to inform their future borough-wide strategy for obesity prevention.

Evaluation of such interventions is problematic due to the usual multi-faceted nature of the interventions and the lack of controlled conditions. Moreover, local authorities are in a state of almost constant flux; programmes, personnel and budgets are continuously under review, and value for money is high on the agenda – all of which bring additional challenges for a commissioned evaluation team to navigate.

In this paper we illustrate these challenges by outlining the approach we took to designing and conducting the evaluation of the pilot programme, using the evaluability assessment approach described by Joseph Wholey and others. An analysis of our approach, using the general principles and key elements of Soft Systems Methodology (Checkland, Poulter 2010), highlights the difficulties we experienced in a number of areas, including: 1) the time and cost implications to develop the right tools to collect the data required, due to the lack of validated and suitable tools to measure dietary or physical activity behaviour change in children; 2) developing a logic model that all parties agreed upon; 3) managing expectations in relation to what outcomes the evaluation team realistically thought could be achieved over the life of the pilot; and 4) retaining independent and academic integrity when being paid by the owner of the programme to evaluate it.

The paper highlights the value of intensive, pre-evaluation activity that helps to develop a pragmatic, shared plan for the evaluation through the process of collaborating with the end users of the evaluation to identify the programme logic and make assumptions explicit. A number of recommendations for others engaging in this type of complex, real-world programme evaluation are presented.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 045 Evaluating Health Programs

## O 141 - Data Use for Decision and Policy Making Health Facilities in Kenya

*E. Karijo*<sup>1</sup>

<sup>1</sup> *Kenyatta University, Public Health, Nairobi, Kenya*

The broad objective of this study was to determine the utilization of the routine generated data for decision making and policy development in health facilities in Kitui County, Kenya. The general objective was achieved by specifically examining the level of utilization of generated routine data, establish the factors influencing utilization of the generated routine data and identifying the influence of the organization of health system on utilization of generated routine data for decision making.

**Methods:** The study was a descriptive cross sectional study adopting Multi stage cluster sampling and simple random sampling methods. The participants interviewed were 110. Data collection also included Key informant interviews, focused group discussions and observation of the available structures in the health facilities. Recordings were transcribed, coded and analyzed thematically. Questionnaire results were also coded, tabulated and analyzed.

**Results:** Overall extent of data use was 66%. The following factors influence use of routine data: The organization of the health system information system influences routine data use such as the number of patients seen daily and the type of the health facility.

The type of records available including those who fill the reports and availability of data tools.

**Conclusion:** The study demonstrated limited utilization of routine data to make decisions for health facility and health sector management.

**Recommendation:** Recommended to the ministry of Health to standardize the parallel reporting levels and unify the reporting tools especially those by different partners. The health facilities in the county level of government should have structured meetings, support supervision and feedback to the health care providers on relevant indicators in the counties.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 046 Evaluating Impacts in Challenging Contexts**

**O 142 - Sound Evidence for Timely Decisions in Volatile Contexts Through Impact Evaluation – Experiences of a Large Implementation Organization**

*F. Krisch<sup>1</sup>, T. Till<sup>1</sup>, M. Florian<sup>1</sup>*

<sup>1</sup> *GIZ, Evaluation Unit, Bonn, Germany*

Within the increasingly dynamic and fragile contexts of development cooperation, providing sound evidence for timely decisions during program implementation has gained importance. Especially in transition phases, project managers need to react fast if the chosen approach does not contribute to the desired impact, whereas conventional evaluations are principally concerned with outcomes of projects and programmes. While traditional monitoring delivers continuous information about achieved outputs, their utilization, and changes regarding outcome indicators over time, the resulting trends do not tell to what extent measured changes can be causally attributed to project activities. Yet, understanding the actual impact is crucial when deciding which mode of implementation to pursue or even to upscale, and which to drop since it does not deliver the expected results.

To enable evidence-based decision-making, an increasing number of GIZ projects and programmes parallel to the implementation of activities are conducting rigorous impact evaluations. These are typically designed to systematically assess the impact of specific interventions. To date, this happens rather decentralized upon the personal initiative of programme leaders who often seek support from university institutes to master the methodological challenges of an impact evaluation using experimental and quasi-experimental designs.

With the aim to start a mutual learning process and to promote organizational standards, GIZ evaluation unit is currently conducting a review of rigorous impact evaluations that accompany project implementation. Based on an internal survey of such impact evaluations using experimental or quasi-experimental design, a document analysis of reports and interviews with programme leaders and partners, as well as evaluators and scientific advisors good practice, major challenges, and untapped potentials shall be elicited. In addition, the review will examine the implementation of a large new impact evaluation of cash-for-work programmes in Iraq to interview relevant stakeholders and learn about success and failure factors. This formative evaluation of an impact evaluation will generate first hand evidence of what works and what not.

At EES, first results of this review plus accompanying research will be presented, focusing on evidence regarding the suitability of accompanying impact evaluations for different project types, the feasibility of different methodological approaches, necessary resources, and the type of (scientific) support needed. Based on the utilization and benefit of the evidence generated, also from the perspective of partners and donors, cost-benefit considerations will be put up for discussion.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 046 Evaluating Impacts in Challenging Contexts**

**O 143 - Managing for Evaluability and Maximizing Use of Evaluation for More Adaptive Programming: Emergent Challenges and Opportunities from Practice**

E. Febles Carmona<sup>1</sup>, S. Khayyo<sup>1</sup>

<sup>1</sup> Oxfam Intermon, International Cooperation Department, Barcelona, Spain

Dealing with development issues in rapidly evolving contexts requires recognizing uncertainty as a key factor influencing all programmes across the programme cycle. This includes uncertainty about how the context will evolve and when, and uncertainty about how change may come about in the specific context. Recognizing this uncertainty and managing our programmes adaptively is one of the challenges confronting evaluation systems and approaches aiming to contribute to more resilient societies. Moreover, as resilience interventions deal with complex social change processes, where resilient solutions are neither straightforward, simple nor clear from the beginning of the intervention, there is a high uncertainty beforehand about the impact of the intervention and most effective routes to building resilience (Guijit, 2017). In this context, evaluation designs and approaches should help programmes to “discover” resilience pathways, providing information that can be used during implementation to adapt interventions, while maximizing evaluability at the end of the project to enlighten understanding about impact and contributing factors. This paper presents the experience of the development and implementation of the evaluation strategy in such context of the EU TF[1] funded resilience project “Employment, resilience and social cohesion in the Sahel strip and Lake Chad area” henceforth *Trust Fund Project*.

The *Trust Fund Project* is a four-year intervention implemented by Oxfam in consortia with CARE, Action contre la Faim (ACF) and three local partners in the regions of Bahr el Gazal, Kanem and Lac, in Chad. The project aims to contribute to resilient development outcomes by strengthening resilience capacities: fostering inclusive local development governance processes, enhancing agricultural and livestock production capacities, and increasing economic opportunities, specifically for youth and women. The project kicked-off in April 2017 and during this first year Oxfam has been leading the design of an evaluation design and implementing baselining activities along with the project team and local stakeholders. The evaluation design proposes a theory-based approach to impact evaluation. This evaluation approach aims to embed a culture of learning and evaluation throughout the project and to improve understanding about how and for whom (intersecting gender and age) project strategies contribute to increasing resilience capacities and thus how this contributes to resilient development outcomes. This evaluation approach acknowledges that bringing in evaluative thinking from the beginning of the intervention plays a critical role in maximizing the use of evaluation findings, improving the quality of data collected and increasing the evaluability of this complex intervention in a changing context.

This paper will present the challenges and opportunities emerging from our experience during this first year on designing and managing this evaluation. We aim to contribute to the debate about the role of evaluation, its design and management in complex interventions implemented in changing environments. This paper will explore how the evaluation design at the onset of the project helped to enhance the quality of the intervention and its Theory of Change, identify evidence gaps, and improve the project's monitoring framework, providing timely and quality data to support adaptive management.

[1] EU TF: European Union Trust Fund for Africa: [https://ec.europa.eu/europeaid/regions/africa/eu-emergency-trust-fund-africa\\_en](https://ec.europa.eu/europeaid/regions/africa/eu-emergency-trust-fund-africa_en)

Wednesday, 3 October 2018  
16:30–18:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 047 Evaluation Capacity Building 1**

**O 144 - An Inverted, Project-based Evaluation Classroom for Evaluation Competence Building**

*L. Balzer*<sup>1</sup>

<sup>1</sup> *Swiss Federal Institute for Vocational Education and Training, Head of Evaluation Unit, Zollikofen, Switzerland*

It is one indicator for professionalization that many basic and further education programs in evaluation have been developed recently. This is good news, because the importance of evaluation in society is unbroken, as is the need for evaluation-specific competencies to conduct high quality evaluation projects which are useful and used. But to teach and learn these competencies is not easy. New and innovative teaching ideas are needed to offer effective trainings to enhance evaluation competence building. In this paper, such an idea with a two-part evaluation training is described.

In the first part of the training, basic evaluative knowledge is being developed. Learning objectives are that participants learn to explain a systematic 10-step evaluation process for planning and implementing an evaluation project, and to apply this knowledge to their own evaluation project. Two didactical principles guide this first part of the training: With the “inverted classroom model” (or flipped classroom; Bergmann & Sams, 2012) the normal procedure of teaching is changed in a sense that what normally takes place inside the classroom (learning knowledge) is flipped with what is normally done outside (transfer/application). Participants learn in advance and outside the phases of attendance, in a self-regulated, asynchronous manner, and with their own learning speed at a location they prefer. This is done on the basis of a textbook which addresses the systematic 10-step evaluation process (Balzer & Beywl, 2015/2018). The second didactical principle is following the “project method” (Frey 2012): Participants are planning a realistic evaluation project within their professional field. During three phases of attendance (1–1.5 days each) nearly no knowledge is taught. Contrary to this, participants deepen their understanding of what they have read and transfer this into their own practical field by creating an evaluation plan for a specific project in their professional context. Teachers change their role and consult participants during this process, which is structured by working sheets which address the 10 steps of the evaluation process. Questions and problems can directly be discussed with real-life examples of the participants. Phases of attendance take place in four week intervals, so in between there is enough time to read the corresponding book-chapters and to deepen the understanding of what has taken place in the last attendance phase. With this procedure, learning is directly linked to transfer into practice. At the end of the first part of the training, all participants have built a plan for a real evaluation project. This is the basis for the second part of the training. If empirical social sciences methods should also be learned, additional courses have to be taken.

In the second part of the training, implementation of the evaluation plans takes place. It is also important to clarify how much empirical knowledge is already available (especially data collection methods and data analysis). Depending on this, participants run their evaluation project with more or less intensive external support.

In this paper, the training concept will be presented as well as first experiences and evaluation data with different implementations.

Thursday, 4 October 2018  
8:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

S 076 Innovative Methodologies and Techniques for Casual Understanding

## O 145 - Reality Bites: Navigating Quality and Use in a Large-Scale Realist Evaluation

I. Vogel<sup>1</sup>, M. Punton<sup>2</sup>, R. Lloyd<sup>3</sup>

<sup>1</sup> Independent Evaluation Consultant, Isabel Vogel Ltd, Haddington, United Kingdom

<sup>2</sup> Itad Ltd, Evaluation, Hove, United Kingdom

<sup>3</sup> Itad Ltd, Organisational Effectiveness, Hove, United Kingdom

This paper reflects on a recently-completed three-year impact evaluation of the £15 million Building Capacity to Use Research Evidence (BCURE) programme, funded by the UK Department for International Development (DFID) and conducted by Itad, UK. The BCURE programme aimed to increase the capacity of policy makers to use research more effectively, through building skills, incentives and systems across 12 countries in Africa and Asia. The BCURE evaluation accompanied the programme over three years.

A realist evaluation approach aimed to explore what worked, for whom, how and in what circumstances, to build capacity for evidence-informed policy making. The evaluation drew on semi-structured interviews with almost 600 stakeholders across 6 countries, as well as programme monitoring data and document reviews.

### Findings and learning points

We found that realist evaluation has the potential to generate practical and nuanced insights that can inform programme design and implementation. However, in attempting to stay true to the principles of realist evaluation, we faced several tensions between quality and usability, including:

- **Quality of analysis vs rapid results for learning.** Realist approaches require significant skill that takes time to build up within an evaluation team, and analysis is potentially very time consuming – which creates challenges when opportunities to contribute to programmes require a quick turnaround.
- **Quality of implementation vs quality of communication.** Like other approaches, realist evaluation comes with its own language and terminology, which creates barriers to the decipherability of findings. It took three years of experimenting for the BCURE evaluation to find a style of presenting findings that was true to the realist approach, but also communicated findings clearly.
- **Whose ideas about quality count?** The evaluation had to navigate tensions between the quality principles of realist evaluation, and donor assumptions about quality that derive from different paradigms. Resolving these tensions was crucial to ensure our findings were viewed as sufficiently robust to inform future decision making.

**References:** The BCURE evaluation report (and other evaluation products including a literature review and practice paper) are available on the Itad website here: <http://www.itad.com/knowledge-and-resources/bcure/>

Wednesday, 3 October 2018  
16:30–18:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 047 Evaluation Capacity Building 1

## O 146 - Experience from the Sri Lanka National Evaluation Capacity Building Programme

*A. Kalugampitiya<sup>1</sup>*

<sup>1</sup> *EvalPartners, EvalPartners, Colombo 5, Sri Lanka*

Parliamentarians Forum for Development Evaluation- South Asia (PFDE-SA) in collaboration with the Sri Lanka Parliamentarians Forum for Evaluation (SLPFE) implemented the project “Strengthening national evaluation system in Sri Lanka” in 2017. This is a continuation of the first project implemented in 2016.

The main purpose of the project is for further step towards strengthening the enabling environment for evaluation through national evaluation policy development, engaging parliamentarians and evaluation capacity building, to promote accountability and evidence based policy making, and facilitate achievement of the SDGs in Sri Lanka. It is the second phase of the Project Management and Evaluation Capacity Development Programmes successfully concluded in 2016 with the support of UNICEF and EvalPartners/EvalGender+ initiative and include further capacity building in 2017 involving a number of partners.

The **objectives** of the project engagement were to:

§ Advocate for National Evaluation Policy (NEP)

- § Develop equity-focused and gender-responsive evaluation guidelines/standards are readily available for reference and use
- § Establish and operationalise Sri Lanka Parliamentarians Forum for Evaluation to advocate for demand and use of evaluation for evidence based policy making
- § Create awareness of the government officials in District Coordinating Committees in 12 targeted districts on SDGs and use of evaluation leading to results-oriented decision-making and effective resource utilisation
- § Enhance capacity of public sector officials to conduct evaluations and monitor SDGs

In the long term these initiatives will facilitate setting up a platform for the legal and policy framework needed to establish a national M&E system in Sri Lanka with a National Performance Management Unit with Integrated Results Based Management (IRBM) at national and sub-national levels.

To implement “Evaluation” agenda two motions were passed in the parliament to formulate a “National Evaluation Policy”. One is formulation of National Evaluation Policy (NEP) and other motion is proposing allocation of funding in the national budget proposal (evaluation budget). It also will help in the process of implementations of an appropriate methodology and allocation of funds from sub-national budget. The Sri Lanka Parliamentarians Forum for Evaluation was also formed in the same year. In year 2016 several evaluation capacity development and SDG awareness-building programmes were conducted to public sector officials at the national and subnational levels as well as members of the civil societies.

In 2017 the programme was continued by Sri Lanka Parliamentarians Forum for Evaluation. Parliamentarians took the lead in organizing the capacity building for sub national level public officials. It was a true example how parliamentarians can play a lead role in promotion of evaluation. Training of Trainers was held to use the evaluation training curriculum developed by the project. The achievements of the project are exciting. This paper is to share the experience of the Sri Lanka National Evaluation Capacity Building Programme.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 048 Enhancing Stakeholder Involvement and Better Learning**

**O 148 - Exit with Learning. A Tale of Three Perspectives:  
Evaluation, Learning and Institutional Development**

*J. Martinho<sup>1</sup>, L. Risby<sup>2</sup>, P. Lacerda<sup>3</sup>, J. Schultz<sup>4</sup>, J. Oliveira<sup>3</sup>, N. Werneck<sup>5</sup>*

<sup>1</sup> C&A Foundation, Effective Philanthropy, Sao Paulo, Brazil

<sup>2</sup> C&A Foundation, Effective Philanthropy, Zug, Switzerland

<sup>3</sup> C&A Foundation, Education & Culture, Rio de Janeiro, Brazil

<sup>4</sup> C&A Foundation, Education & Culture, Sao Paulo, Brazil

<sup>5</sup> Independent Consultant, Not Applicable, Rio de Janeiro, Brazil

C&A Foundation is a corporate philanthropic entity based in Switzerland and its primary emphasis is on four programmes focused in the Fashion Industry: Sustainable Cotton; Working Conditions; Forced and Child Labour; and Circular Economy. In 2013 it was re-structured to lead and manage all of the corporate philanthropic entities of the C&A company in retail and sourcing countries, including Fundación C&A in Mexico, and Instituto C&A in Brazil. Since 1991, when it was created, the Instituto C&A funded mainly Education and Culture projects and was recognized as one of the main funders in Brazil. For this reason, during the restructuring process, Instituto C&A had the opportunity to plan a three-year phase out of the education portfolio. The purpose? Exit with Learning. To achieve this, Instituto C&A developed and executed an ambitious learning agenda through evaluation focused on three main objectives: systematizing what we have learned to support future C&A Foundation initiatives, structuring the legacy we intend to leave to the Brazilian civil society and other funders, and supporting the institutional development of funded partners. Still, in addition to the natural challenge that a daring agenda like this poses, Brazil has undergone a marked economic and political crisis that has challenged both the agendas advocated by the partners and their own sustainability. For this reason, Instituto C&A chose to rethink the focus of two external evaluations it had planned and, from evaluations designed by the funder and focused mainly on outcomes and learning, they went on to co-create both external evaluations with the partners. As one would imagine, this approach eventually brought several challenges that a more traditional approach would not bring, but also made it possible to achieve results that would otherwise not be achieved. During this presentation, we will discuss the challenges, results and what we have learned from this process. We will try to argue the ability of an evaluation to be more than a mere validation of achievement of outcomes or a systematization of learning. In this sense, we will debate how an evaluation may also be a powerful tool for institutional development and particularly useful in times of crisis.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 048 Enhancing Stakeholder Involvement and Better Learning

## O 149 - Evaluating Community Ownership: The Advantages of the Feminist Standpoint Approach

*P. Banerjee*<sup>1</sup>

<sup>1</sup> DIKSHA Non-governmental social service organisation, Co-founder, Kolkata, India

**Rationale:** The intended paper will be based on actual experiences of evaluations conducted by the author, in each of which 'community ownership' was an important criterion to be assessed. The challenges faced in that context were complex enough to raise questions about the very understanding of 'community ownership', which constitutes one of the key focal points for evaluating many social development interventions, especially from the perspective of sustainability. It is important, therefore, to raise these complexities and the lessons learnt from them, to facilitate collective rethinking of both the notion of 'community ownership' and of methodologies best suited to measure the degree and quality of such ownership.

#### **Objectives:**

1. Through actual case-studies, demonstrate why 'community ownership' is a tricky and complex notion in itself'
2. Present an overview of methods used to address those complexities for an in-depth evaluation using participatory methods.
3. Showcase why 'objective' evaluation by a 'distanced' assessor may well be inadequate in such complex situations and argue in favour of the strengths of the Feminist

Standpoint approach.

**The Paper:** 'Community Ownership' in any social development intervention is an important factor to be evaluated from the perspective of sustainability, as the Sustainable Development Goals amply demonstrate. But 'community ownership' is a tricky phenomenon since it always involves an aspect of representation from the community. Who decides the criteria of such representation is one important question. What should be the evaluator's analysis when such representation leaves out one or more sub-groups within a community outside the representation? Discussions around these two important questions will be the pivot for this paper.

Through three different case-studies, it will be demonstrated that until and unless a very fine gender lens is used to assess the degree and quality of 'community ownership' – such intricate details may lie beyond the evaluator's notice. The gender lens in this context does not refer merely to a critical eye for the inclusion of girls and women, but also to a meta level of understanding about the power dynamics within a community and the inclusion of those lower down on that power scale. It will be argued through these case-studies that a 'distanced' evaluator keen only on maintaining the so-called objectivity may well miss these nuances. The feminist standpoint approach, which allows the assessor to connect to the issues at hand with her full subjectivity, is far more suitable for understanding these intricacies. Such involvement does not diminish the value of the evaluation; rather enriches it substantially by putting the human subject in the centre.

The main argument of this paper will be to provoke a serious methodological rethink on the so-called subjective/objective divide by pointing out the deficiencies of 'objective scientificity' – which obfuscates subtle nuances of lived-in experiences. Since the case-studies will be located in different developing and under-developed countries – there will be scope within the narrative to argue that the feminist standpoint approach is suitable across cultures and population groups and needs to be taken far more seriously in evaluations than it currently is.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 049 Humanitarian/ Development Nexus**

## **O 150 - The Humanitarian-Development Nexus- What Do Evaluations Have to Say?**

*M. Bruno<sup>1</sup>, F. Bonino<sup>2</sup>, A.C. Luzot<sup>3</sup>*

<sup>1</sup> UN Food and Agriculture Organization, Office of Evaluation, Rome, Italy

<sup>2</sup> UN High Commission for Refugees, Evaluation Service, Geneva, Switzerland

<sup>3</sup> World Health Organization, Evaluation Office, Geneva, Switzerland

This paper represents 'work in progress' of the UNEG Humanitarian Evaluation Interest Group. At its origin, is a question on whether and how evaluations have covered the interface, or nexus, between humanitarian and development work and, if so, what exactly have they covered and how? The question directly relates to the need, expressed in terms of commitment, in the Agenda 2030 and the World Humanitarian Summit, of delivering on collective outcomes and transcending the humanitarian-development divide. The assumption is that evaluations can contribute to a better understanding of the lessons and emerging issues through a cross-cutting analysis of evaluative findings and practices. The study reviewed a purposive sample of 123 evaluations, of which 97 focused on specific countries and 26 were global.

The mapping is exploratory in nature and tries to identify possible patterns across different evaluation commissioners, type of evaluations, scope and focus, as well as the specificities of nexus evidence and narratives in each of the countries covered as case studies. The report also records gaps and good or promising evaluation practices that contributed to a better understanding of the humanitarian-development interface. A lack of shared definitions of the nexus and its broad and evolving conceptual boundaries meant that the criteria for inclusion and exclusion of reports were discretionary, making it also too complex to identify a statistically significant universe. Evaluations selected include a host of issues expected to relate to the nexus, such as agencies' policies and performance related to disaster risk reduction, resilience, fragile, crisis and conflict contexts, evolving delivery modalities and programming coherence, financing mechanisms and inter-agency coordination. As the findings were emerging, the study team conducted interviews with 10 key informants to investigate the expectations and demands around the use of evaluative evidence.

The specific research questions used, were:

- To what extent and how have humanitarian and development evaluations considered the topic of the nexus? Can significant differences be observed in how evaluations have covered the topic of the nexus?
- Which type of evaluative evidence has been generated about the humanitarian-development nexus?
- Does a preliminary mapping and description of the evaluative evidence from the sample selected for the study point to significant gaps, good practices and lessons?

The objective of the paper is to present this preliminary mapping and synthesis as a conversation starter amongst different stakeholders, evaluators and researchers faced by the challenge of evaluating nexus-related areas, as well as to contribute to the learning and progress made towards international commitments such as WHS and SDGs.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 049 Humanitarian/ Development Nexus**

**O 151 - Building Bridges Between International Humanitarian and Development Responses to Forced Migration**

R. Wedel<sup>1</sup>, A. Kocks<sup>1</sup>, H. Roggemann<sup>1</sup>, H. Roxin<sup>1</sup>

<sup>1</sup> DEval, Governance- Bi and Multilateral Development Cooperation, Bonn, Germany

The importance of bridging humanitarian assistance and long-term development cooperation is a recurrent theme in the political debate on how to deal effectively with protracted forced migration crises. One common denominator in the debate is that such protracted crises can be dealt with by creating conditions that lead to long-term improvements and increased resilience for all affected groups: refugees, IDPs and host populations. In this context, the 2016 World Humanitarian Summit reached a consensus of the need to better link humanitarian assistance and development cooperation. But the agreement leaves open how this can best be done in practice. Moreover, there is not much robust evidence and knowledge about the extent to which effective linkages have already been established in practice. To date there is hardly any evaluation that explicitly examines such linkages in the context of forced migration crises.

The present joint literature review of the German Institute for Development Evaluation (DEval) and the Swedish Expert Group for Aid Studies (EBA) presents an overview of knowledge on the humanitarian-development gap and how it could be closed:

In a first step, we systematically analyzed concepts of the humanitarian-development nexus with regard to how they characterize the humanitarian-development gap from their individual perspective, and how they propose to bridge it. Grouping these characterizations and recommendations according to analytical categories, then allowed us to describe the abstract concept of the humanitarian-development gap in a more structured and concrete manner along different dimensions (i.e. sub-gaps) of the main gap. Overall, we identified seven different dimensions of the humanitarian-development gap and nineteen recommendations filtered out of the conceptual sample studies on how to close the gap with respect to these dimensions.

The findings are organized in an *innovative analytical framework*, allowing for a structured linkage perspective on individual forced migration crises. It represents a leap forward to conceptual clarity as it captures the state-of-the-art knowledge on the linkage discourse by synthesizing the most relevant concepts into a single analytical framework. The application of the analytical framework enables evaluators to identify elements that promote or hinder the linkage of humanitarian assistance with development cooperation.

In a second step, we utilized this framework to analyse evaluations and evaluative studies on the international response to the Syria crisis. The framework enabled us to explore to what extent (and why or why not) the conceptual recommendations are translated into actual deeds. The findings suggest some achievements in linking humanitarian assistance with development cooperation after the early years of the Syria emergency, when all focus was placed on response to crisis management. Yet, findings expose a lack of coordination and show to policy makers and practitioners what still needs to be done in addressing remaining sub-gaps.

Future evaluative work can build upon our conceptual and empirical findings by developing a proper theory of linkage (a Theory of Change, ToC), one which spells out exactly how to link what at various levels, what the intended effects of these linkages are, and exactly how to achieve these effects.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 050 Evaluating Livelihood and Community Programs**

**O 152 - Impact Evaluation of Interventions for Climate and Livelihood Resilience: Lessons from a Livestock Insurance Scheme in Ethiopia**

*M. Meller<sup>1</sup>, M. Frölich<sup>2</sup>, G. Montresor<sup>3</sup>, E. Ayifah<sup>2</sup>*

<sup>1</sup> *Particip GmbH, Evaluation Unit, Freiburg, Germany*

<sup>2</sup> *University of Mannheim - Faculty of Economics, Chair of Econometrics, Mannheim, Germany*

<sup>3</sup> *University of Mannheim, Center for Evaluation and Development C4ED, Mannheim, Germany*

The current study evaluates the impact of the “Satellite Index Insurance for Pastoralists in Ethiopia (SIPE)” Programme, an ongoing pilot intervention sponsored by the World Food Programme. SIPE seeks to strengthen the resilience of pastoralist households to drought shocks by offering a livestock insurance scheme. Insurance payouts to households are triggered by droughts identified through satellite images. The impact evaluation design for this study departs from a randomised controlled trial (RCT) at household level.

The resilience aspects of the intervention, as well as the geographic context of high vulnerability to climate shocks, create several challenges for the impact evaluation approach. This presentation highlights these challenges (including ethical issues) and proposes solutions, which may apply to other impact evaluations in similar settings.

The first challenge is the fact that pastoralists within the same villages of the programme region have a strong sharing culture. Consequently, coping strategies for climate shocks and potential insurance benefits typically extend across households. These spillovers may potentially bias impact estimates if – for reasons of programme design or implementation – randomisation of beneficiary households is only feasible within the selected pilot villages. This study adopts an innovative evaluation design that involves a second control group of pastoralists in non-pilot villages. Impact estimates are adjusted by estimating the extent of spillovers in two steps. Programme beneficiaries in pilot villages are first compared against non-beneficiaries in SIPE villages (by RCT), and then against non-beneficiaries in non-pilot villages (by propensity score matching).

Second, the evaluation design has to deal with the unpredictability of drought periods insured by the intervention. Since it is uncertain whether a drought will actually occur and trigger insurance payouts in the study period, the evaluation has been designed to address both scenarios with and without drought. Evaluation questions, result indicators and data collection tools have been developed to account for both scenarios under unpredictable climate shocks.

The third challenge, with implications for data collection, stems from the semi-nomadic grazing patterns of livestock holders in the programme region. Pastoralists move around in line with seasonal rainfall patterns and climate shocks. This may complicate the traceability of baseline respondents for longitudinal data collection in the endline survey. Potential solutions consist in scheduling the survey in periods with reduced mobility, collection of migration patterns and cell phone numbers in the baseline, and accounting for attrition in the sampling strategy.

Friday, 5 October 2018  
11:15–12:45

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 128 Promoting Use of Evidence in Government**

**O 153 - Large-Scale Global Surveys: Approaches for Data Collection and Validation for a Family Planning Policy Study**

*S. Gold<sup>1</sup>*

<sup>1</sup> *Chemonics International, USAID Global Health Supply Chain Program - Procurement and Supply Management GHSC-PSM project, Washington- DC, USA*

How do you conduct a survey of public health officials in nearly 40 countries concerning family planning policy and finance, and ensure that the data is accurate? The United States Agency for International Development (USAID) Global Health Supply Chain Program – Procurement and Supply Management (GHSC-PSM) project based outside of Washington, D.C. (USA) set about to accomplish this task, with the goal of increasing the body of knowledge in family planning programming and advocacy in developing countries. The research was structured to provide a snapshot of the existing landscape, as well as detect potential systemic attributes that may facilitate, or hinder, contraceptive security. Our approach also incorporates activities designed to improve data use, by providing stakeholders with opportunities to engage with the researchers, providing follow-up, and disseminating our results through different channels.

The study builds on an effort implemented by the USAID DELIVER project since 2009 to collect data on contraceptive security, the condition where everyone is able to choose, obtain, and use high quality contraceptives when they need them for family planning and the prevention of sexually transmitted diseases. This paper will elaborate on methods for survey development, and remote data collection and validation of a wide-scale global survey, where the respondents are primarily government officials in resource-poor settings. The paper will discuss approaches to maximize the response rate and data validity under the constraints of administering the survey remotely.

In a multi-step process, GHSC-PSM employed several strategies to reach out to governments and donor agencies to elicit a response, including in many countries where the project presently has no resources. Involvement of USAID officials both in Washington and USAID missions in countries was key, as was connecting with other donor representatives. GHSC-PSM further conducted both a landscape analysis to ensure the uniqueness and relevance of each data point to be collected, and a market analysis to better understand how current and potential users preferred to access and use the data. The survey was pilot-tested by in-country project monitoring and evaluation teams.

The survey comprises both quantitative and qualitative elements and is populated through key informant interviews mainly of government officials. In the past, data came from numerous, mostly unspecified sources. GHSC-PSM refined several key steps along the process from survey development to data validation to standardize both the sources of data and the methods for validation. Key strategies to be discussed include the inclusion of a data collection and usage manual in four languages, limiting responses through drop-down menus, requiring respondents to select from a specified and pre-tested menu of data sources for each question, use of a detailed validation checklist, and triangulation with other primary and secondary data sources.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 050 Evaluating Livelihood and Community Programs****O 154 - Comparison of Livelihood Resources Among Beneficiaries of Critical Ecosystem Management Project in Ajasse-Ipo and Koton-Karfe Watersheds in Nigeria***M.O. Olaolu<sup>1</sup>, E.A. Agwu<sup>2</sup>, M.C. Madukwe<sup>2</sup>*<sup>1</sup> *University Of Nigeria- Nsukka, Department Of Agricultural Extension, Nsukka, Nigeria*<sup>2</sup> *University of Nigeria, Agricultural Extension, Nsukka, Nigeria*

The Second National Fadama Critical Ecosystem Management Project (Fadama II- CEMP) operated on the broad objective of maintaining the productive and ecological health of the Fadama resources base- watersheds, in order to enhance the productivity of the Fadama areas and the livelihood systems they support, through sustainable land-use and water management. The project was implemented in five different watershed areas in Nigeria. But the 2012 flood incidence in Nigeria had different impact on these watersheds during the project implementation with Koton-Karfe having a greater impact than that of Ajasse-Ipo. This was in spite of the fact that the two watersheds are within the same area. This study sought to ascertain the interference of the flood incidence on the expected impact of the project on beneficiaries' livelihood resources. Resources, capabilities and activities required for a means of living is the essence of a livelihood, this study therefore, aims at comparing the project beneficiaries' with respect to beneficiaries' natural livelihood resources; financial (liquid and earned) livelihood resources and social livelihood resources in two watersheds among five watersheds that participated in the intervention project. The study was carried out in Kogi (Koton-Karfe watershed) and Kwara States (Ajasse-Ipo watershed) of Nigeria using survey design. A multi- stage sampling technique was used to sample respondents 120 beneficiary farmers (60 each from the watersheds) but 54 and 60 beneficiaries in Kwara and Kogi States respectively were eventually interviewed giving a total of 114 beneficiaries. Interview schedule was used to collect data for the study. Data was presented using frequency, percentages and mean while student t-test was used to compare these two watersheds. The study found that there is a significant difference on mean number of economic trees on their lands ( $t=11.8$ ,  $P<0.05$ ) in favour of Ajasse-Ipo watershed. There is also a significant difference in the bags of soybean (100kg) ( $t=2.2$ ,  $P<0.05$ ) and rice (50kg) ( $t=-3.9$ ,  $P<0.05$ ) owned by these beneficiaries in the two watersheds. With respect to beneficiaries' social livelihood resources, the two watersheds do not have any significant differences in the mean number of labour based/ exchange groups (formal) ( $t=-0.9$ ,  $P>0.05$ ) and market assisting groups ( $t=0.0$ ,  $P>0.05$ ). Beneficiaries from Ajasse-Ipo watershed earned an average of 150,000 Nigerian Naira while those of Koton-Karfe watershed earned nothing from remittances from natural endowments. The study found that though the two watersheds benefited from the project at the same time and with similar baseline condition, the Ajasse-Ipo watershed had higher impact on their natural and financial livelihood resources as against that of the Koton-Karfe watershed from the project. This is because the Ajasse-Ipo watershed had far better improvement in these livelihood areas over the Koton-Karfe watershed though in other areas the two watersheds were still at the same levels. The study therefore, recommends that agricultural interventions should also put into consideration the vulnerability/ proneness of different project sites to natural and man-made disasters for possible adjustment in the magnitude and approach of intervention. This will be possible if only interventions are flexible in the cause of implementation.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 051 Evaluation and Government Transparency

## O 155 - Re-Examining Evaluation Design to embrace the Sustainable Development Goals: Lessons and Best-fit Practices from Multilateral Organizations

C. Tarazona<sup>1</sup>, A. Ould Abdallahi<sup>1</sup>

<sup>1</sup> FAO, Office of Evaluation, Rome, Italy

**Rationale:** In recent years, and following an era of unprecedented economic growth and globalization, many developing countries and their citizens have become more engaged and vocal about accountability and results. In many countries, eradication of poverty, malnutrition and hunger have become visible national goals, and thus national policies and programmes have been developed to pursue them.

In parallel, several governments have embarked in developing/strengthening their national monitoring and evaluation systems. Designing and directing development policies and humanitarian actions are however not solely under the control of government technocrats and/or development entities. There is a much broader participation of civil society, academia and citizens themselves in monitoring, debating and criticizing public actions, both nationally and internationally. The inclusive consultation process which led to the adoption of the Sustainable Development Goals (SDG) and the Agenda for Humanity exemplify this trend.

A major lesson from the SDG process is indeed that formulation of development actions is more credible if it integrates the views of a broad set of stakeholders, which in turn is based on a proper contextual analysis. Likewise, the evaluation of SDGs should be based on a similar process, and be undertaken in such a way that the results inform national planning/monitoring processes. For the UN and multilateral organizations, the inclusion of national and local actors who are responsible for driving development process towards the achievement of the SDGs is a must. In order for their evaluations to contribute to the attainment of these goals, multilaterals must devise ways to reach out to these national and local actors, who are the potential users and beneficiaries of these evaluations, and tailor the design of the evaluations so that they capture the key aspects, contributions and lessons from developmental interventions.

**Objectives sought:** This paper will discuss the experience of selected United Nations and Multilateral Development Banks (MDBs) to take into account the 2030 agenda, especially in terms of evaluation inclusiveness and criteria. Recent attempts to strengthen inclusiveness and develop tailored evaluation criteria in strategic and country level evaluations of the FAO will be used as primary examples. In addition to FAO's examples, this paper will rely on the following analyses:

- Stock-taking of articles, papers and discussions taken place within the UN Evaluation Group on SDGs and partnerships.
- A review of corporate evaluation strategies/plans within the UN as well as of a sample of large evaluations conducted by major UN Agencies and MDBs post-SDGs to analyze SDGs consideration in evaluation plans and designs.
- A survey of evaluation heads and managers from selected entities on readiness for SDG contribution assessments.

The paper will consist of three main sections: i) introduction to SDGs and mapping of current practices at FAO and in selected multilaterals (state of affairs); ii) promising steps and ongoing initiatives to include SDGs-related considerations into corporate evaluations (state of the art), iii) ways forward for multilaterals and other development actors. It is hoped that this research will be of interest not only to the EES but also the wider development evaluation community including the UN and the MDBs.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 051 Evaluation and Government Transparency

## O 156 - Learning Partnerships: Successfully Generating Purpose-Driven Monitoring and Evaluation (and How to Strike the Balance Between Flexibility, Collaboration and Independence)

C. Smit<sup>1</sup>

<sup>1</sup> Genesis Analytics, Evaluation for Development, Johannesburg, South Africa

In a context of greater scrutiny of government and donor spending, rapidly increasing social impact investment, and the need to maximise the impact of these development efforts, the need for monitoring, evaluation and learning (MEL) has never been higher. Despite this, MEL is sometimes seen as a grudge purchase, time consuming and burdensome for programme staff or a tick-box exercise separate from core activities. Learning partnerships offer a solution to this – a learning partner facilitates meaningful engagement with decision makers to generate the insight they need to improve their interventions and achieve better results.

This presentation will examine:

- a) What is a learning partner? The scope of a learning partner is determined by the needs of the partner organisation and can cover a wide range of MEL tasks. Importantly, a learning partner is embedded in the full lifecycle of a programme or intervention – from design through to implementation and close-out.
- b) The challenges with learning partnerships – learning partnerships aren't always possible, they require resources and a willing and committed partner organisation. Being embedded in a programme and helping to further its goals means that a balance needs to be struck by the learning partner between *collaboration* and *independence*.
- c) How to structure learning partnerships – successful partnerships are developmental, rigorous, flexible, responsive and promote accountability for learning
- d) How to drive greater evidence use – understanding of the partner's context, combining technical and sectoral understanding and being able to prioritise lessons and recommendations all play an important role.

Each section will draw on Genesis Analytics' experience from four learning partnerships: 1) an initiative aimed at catalysing new, sustainable employment opportunities and skills training for African youth, with a focus on the ICT sector, funded by a large global foundation, 2) a 4-year project aimed at scaling impact in client protection for low income financial service customers in Rwanda using a market systems approach, 3) a South African foundation's portfolio of consumer financial education programmes and 4) a large African foundation aimed at providing support to children and young people, from early childhood development, to basic education and bridging programmes.

This presentation aims to propose learning partnerships as a means of facilitating increased engagement by decision makers with insight and evidence, ultimately improving the resilience, inclusion and productivity of communities reached by policies and programmes.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 051 Evaluation and Government Transparency

## O 157 - From Co-Creation Towards Co-Evaluation: Innovation in Evaluating the Effectiveness of Public Policy Development and Implementation in OGP Action Plans

A. Routzouni<sup>1</sup>, A. Deligiannis<sup>2</sup>

<sup>1</sup> University of the Aegean, Department of Information & Communication Systems Engineering, Samos, Greece

<sup>2</sup> International Hellenic University, School of Science & Technology, Thessaloniki, Greece

Open and participative governance is quickly becoming one of the main policy priorities to modernize and reform public administration at a global level. Formulating and promoting public policies for empowering citizens, enhancing transparency, promoting integrity and accountability are central to government reform efforts.

One of the most notable initiatives to support the global open government agenda is the Open Government Partnership (OGP), a multi-stakeholder initiative launched in 2011. It brings together over 75 countries and 15 subnational governments that have committed to making their governments more open and accountable. Every two years, each participating country delivers a National Action Plan (NAP) which is a product of a co-creation effort between the government and civil society to develop concrete, time-bound and measurable open government commitments.

Recognizing that the process of formulating open government policies and creating OGP NAPs involves a wide network of stakeholders and policymakers, the paper studies the continuous involvement of this network in the evaluation of the NAPs quality and the effectiveness of commitment implementation.

OGP engages in a multi-stakeholder co-creation and co-evaluation process since it believes that collaboration between governments and civil society organizations can enhance the effectiveness of public policies and reform evaluation. Including stakeholders and interested parties in the evaluation process can significantly enhance the effectiveness of the process.

The paper examines the process, methods and criteria used by OGP to evaluate NAPs and defines the key aspects, challenges and shortcomings of the collaborative processes used by OGP to evaluate the effectiveness of public policy implementation.

The study analyzes the OGP multistakeholder evaluation approach and the methods used to further engage government officials, civil society and other stakeholders that are involved in the co-creation process and their subsequent involvement in the evaluation of NAP implementation.

Using National Action Plan text, OGP assessment reports and external reviews the study analyzes the process and criteria used by OGP to evaluate the effectiveness of the multistakeholder process used for the co-creation of NAPs in each country. It also identifies and examines the processes and criteria used by OGP to evaluate the quality of the commitments included in the NAPs in each country and those used to evaluate NAP commitment implementation.

By examining the effectiveness of the process to evaluate the development and implementation of public policies the study provides insights into the best practices and key challenges of collaborative evaluation processes. It defines the key characteristics, parameters and priorities of a collaborative evaluation process in order to propose an innovative, open and inclusive assessment process.

The study describes the mechanisms that enable all parties to jointly own and develop the process instead of merely consulting with the government on developing an OGP NAP. It also illustrates the key challenges in the OGP co-evaluation process in order to empower government officials, civil society, and other stakeholders to stay involved beyond the co-creation phase.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 061 Participatory and Stakeholder Approaches

## O 158 - Simple, Participatory Assessment of Real Change (SPARC): Lessons in Adapting Outcome Harvesting for an HIV Prevention Advocacy Network in Africa

*J. Dasmariñas*<sup>1</sup>

<sup>1</sup> AVAC, Policy and Data Analytics, New York City, USA

This paper presents insights generated from a participatory evaluation and learning process called SPARC (Simple, Participatory Assessment of Real Change) that AVAC, a New York City-based international NGO, implemented for its HIV prevention advocacy program in Africa beginning in 2017. By reviewing the seminal experience, the presentation seeks to unpack the potentials and implications for undertaking a complexity-aware approach for an entire advocacy network as well as contribute to enriching the discourse on complexity evaluation.

Following the observation of Michael Quinn Patton and Ricardo Wilson-Grau about the growing demand for creative evaluation approaches that assess complex interventions, SPARC was created to provide a dialogic space for surfacing outcomes in the HIV prevention advocacy field. Drawn from outcome harvesting but modified to adapt to the peculiarities of AVAC's intervention, SPARC was designed to engage a network of stakeholders with varying levels of M&E capacity in identifying signs of progress that go beyond outputs and targeted results.

Infused with reflective exercises where art and games become de facto rehearsals for narrating stories of change, SPARC uses these tactics as foundation for staging the creative potential of program evaluation. In so doing, program partners become evaluators who interactively collect, analyze, and make sense of outcomes.

SPARC as a technique gravitates from organizational engagement to network collaboration to understand how the whole is more than the sum of its parts. This requires a process where individual core partners sift through the most significant outcomes for the year – positive, negative, manifest, latent – through facilitated exercises and moderated discussions. Organizational storylines are then drafted and validated, and eventually presented at the annual SPARCFest where all partners convene to engage in collective analysis and identification of the most significant results through outcome clustering, debate, and consensus-building.

There are certain practical implications uncovered during the initial implementation of SPARC. First, the approach requires a sensitivity to the diversity of network actors and their work streams, a challenge not typically encountered if the evaluand were just a single organization. Second, using internal reporting forums can be used as a foundation for SPARC stories but needs to be unfettered from planned activities. Third, simplified evaluation techniques do not necessarily translate to trading-off robustness of the offered results. Kernels of substantiation and evidence-building can still be integrated through careful guidance. Fourth, sustaining SPARC means promoting partners' independent use and adaptation of the approach and creating champions along the way. Fifth, attaching SPARC to a workplanning process elevates its value for participants wherein collated outcomes become prompts for developing future program actions. And sixth, facilitating the cognitive shift from outputs to outcomes requires a continuous mentorship that ultimately helps reorient partners towards good evaluative thinking.

The presentation argues lastly that a complexity-aware evaluation approach such as SPARC, albeit the diametric opposite of a logic model, can be used as a fertile and nuanced complementary to interrogating results beyond targets. In surveying the field of outcomes writ large, SPARC can prove to be a meaningful learning experience that transcends "check-the-box M&E" and recuperates the dynamism of evaluation practice.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

## S 052 Evaluation of Funding Schemes for Knowledge Production

**O 160 - Rethinking Impact Measurement of University Development Cooperation – Strategies for Modular Evaluation Designs***O. Almqvist<sup>1</sup>, L. Raetzell<sup>1</sup>, F. De Maesschalck<sup>2</sup>*<sup>1</sup> Syspons GmbH, Evaluation Department, Berlin, Germany<sup>2</sup> Belgian Ministry of Foreign Affairs, Special Evaluation Service of Belgian Development Cooperation SEO, Brussels, Belgium

**Impacts** of university development cooperation have rarely been analyzed on the basis of robust evaluation designs, even though **impact orientation** has become increasingly important in this field. Most evaluations are based on a qualitative approach, which is seldom supplemented with multivariate quantitative or robust qualitative methods.[1] One reason for this is that university cooperation often aims for changes at the organizational level through capacity development measures. In addition, often **different definitions of impact** are used within university development cooperation.[2]

Given the central role of academic training and research to foster **resilient societies** in developing countries, **addressing knowledge gaps and methodological challenges** regarding robust evaluations in this field is crucial. In 2017 / 2018, an **impact evaluation of the entire Belgian university development cooperation** in the years 2000 – 2016 has been carried out. This evaluation had a formative and a summative objective. The formative aspect of the evaluation analyzed the definitions of impact used by stakeholders in the field, and assessed the **evaluability** of interventions on the basis of an online and a Delphi-survey. For the summative aspect of the evaluation, a model for the measurement of capacity development in universities has been developed. Following the approach of Stern et al.[3], a **combination of different evaluation designs** for impact measurement (experimental, generative / mechanistic, etc.) was applied in a modular approach and tested for feasibility in different contexts. This included the evaluation of scholarship schemes on the basis of a stratified cohort with a comparison group. With this approach, the longitudinal development of scholarship recipients could be captured retroactively without implementing a panel survey. In addition, the impact of institutional cooperation between universities and university departments was analyzed with quasi-experimental designs, most significant change approaches and before-and-after comparisons. The case studies included **challenging contexts for data collection** because of the lack of a baseline for projects which started more than a decade ago. In addition, the case studies relied on cooperation and intrinsic motivation from stakeholders in Belgium, Benin, Ethiopia and Vietnam whose projects had ended several years ago.

The findings related to the formative objective of the evaluation are used to **increase the evaluability** of Belgian university development cooperation by **rethinking evaluation methods, design and criteria** and identifying tailor-made evaluation designs for robust impact measurement. Findings related to the summative objective of the evaluation are used to increase the impact of Belgian university cooperation for resilient societies (e.g. with regard to brain drain and gender).

[1] Mawer, M. (2014). A study of research methodology used in evaluations of international scholarship schemes for higher education. London: Commonwealth Scholarship Commission in the UK.

[2] Walters, H. (2007). Capacity Development, Institutional Change and Theory of Change: What do we mean and where are the linkages – A Conceptual background paper. Wageningen: Wageningen Universiteit.

Hearn, S. & Buffardi, A. L. (2016). 'What is impact?'. A Methods Lab publication. London: Overseas Development Institute.

[3] Stern, E. et al. (2012). Broadening the range of designs and methods for impact evaluations. Working Paper 38. London: Department for International Development.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 053 Challenging traditional approaches to evaluation**

## **O 161 - Progressive Evaluation: Challenging Results Driven Culture**

*A. Guitard<sup>1</sup>*

<sup>1</sup> *Ian C Davies Conseil, Evaluation, Paris, France*

**Rationale:** “Mainstream” or “traditional” evaluation, as it is institutionalised in the field of international development, remains firmly anchored in an accountability for results paradigm. Although evaluations commonly state a dual purpose of accountability and learning, their implicit causal approach and related suite of methodologies, show that there is still progress to be made to reap fully the value of evaluation.

**Objectives sought:** The predominance of the accountability paradigm in “traditional” evaluations tends to reflect organisational cultures that are often hierarchical, bureaucratic, risk-averse and compliance oriented. In these contexts “traditional” evaluation remains relevant, whereas less so “non-traditional” learning driven evaluations. This presentation explores practical ways in which learning driven evaluation, i.e. progressive evaluation, can be introduced in these contexts.

**Brief narrative and justification:** “Traditional” evaluation tends to focus on the achievement of results and production of an independent report that provides accountability information on the extent to which these have been achieved or not. The report contains conclusions, recommendations and, increasingly, lessons learned. Most of the use of the evaluation report involves management responses to the recommendations and, sometimes elaborate and costly, systems of follow up and reporting on the implementation of recommendations. The report itself is usually subject to external quality assurance. In comparison, very little effort appears directed to the learning component of evaluation.

A number of entry points or levers appear promising to begin to shift organisations towards progressive evaluation. These may be the organisation’s evaluation function itself so that it conceives and formulates its value proposition as one that generates evidence and knowledge to support management and policy, i.e. where the evaluation function is part of the intelligence of the organisation.

Other ways are to implement progressively evaluations that use diverse approaches, i.e. bases for valuing, such as systems referenced, gender based, rights based, theory-driven, etc.; as well as evaluations that are more functional in nature, such as developmental evaluation, ex-ante evaluations, etc. Finally, working with, but distinctly from, more accountability focussed function such as audit and control, can help the organisation to see and understand better the specific and distinct value of evaluation.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 053 Challenging traditional approaches to evaluation**

**O 162 - Questioning the Relevance of Mainstream Evaluation Approaches in a Complex Environment: An African Story**

*J. Toulemonde<sup>1</sup>, P. Jackson<sup>2</sup>, S. Hachem<sup>3</sup>*

<sup>1</sup> *Independent Expert, I, Journans, France*

<sup>2</sup> *African Development Bank, Independent Development Evaluation, Abidjan, Ivory Coast*

<sup>3</sup> *Multilateral Organisation Performance Assessment Network MOPAN, Paris, France*

Our presentation will question mainstream evaluation approaches as far as development results unfold in unpredictable contexts. The discussion will be rooted in a real life story by looking at four levels of connected evaluation works: field level interventions, country strategies, donor organisation as a whole, and multilateral assessment of that organisation. We will identify areas of unpredictability in these evaluation works and reflect upon the unintended negative consequences of mainstream evaluation approaches in such unpredictable areas.

Development aid evaluations tend to apply three mainstream approaches: logical framework, performance ratings, and DAC evaluation criteria. All these three approaches push towards setting objectives that are easily evaluable, i.e. short term, intervention-level, measurable, predictable, and attributable. Evaluators use these narrow lenses for assessing performance in the field. Then field level performance ratings are aggregated at upper levels (e.g. sector, country, donor organisation) in order to satisfy donors' accountability requirements. This is questionable.

When shifting from field level interventions to higher level strategies, evaluators should also shift from specific projects to wider works (e.g. knowledge broking, policy advice), from hard to soft achievements, from short-term and micro outcomes to mid-term and macro ones, and from early to further steps in the causal chains. Evaluators of higher level strategies should focus on what is called impact and sustainability at the level of field interventions instead of crunching performance numbers that refer to narrowly designed objectives. Conversely, they should recognise the influence of multiple players interacting through numerous feed-back loops within networks and systems that are far from stable and loosely circumscribed. They should enter into the world of unpredictability instead of remaining trapped into the implicit assumption that outcomes are predictable enough for setting evaluable objectives.

When donor organisations strive to influence changes in unpredictable contexts, they should be accountable for their agility and responsiveness while mainstream evaluation approaches tend to frame their performance in achieving predetermined objectives.

Our presentation will discuss the above quoted concerns in the context of a real life story which focuses on African Development Bank (AfDB) mainly. We will look at a chain of evaluation works that cuts across several levels of management: [1] assessments of AfDB supported projects (e.g. private sector programme and budget support) in stable and fragile contexts, [2] evaluations of AfDB country strategies and programmes, [3] AfDB's comprehensive evaluation of development results, and [4] MOPAN's assessment of multilateral organization, including that of AfDB.

We will first discuss the extent to which mainstream evaluation approaches applied to the reviewed evaluations and why. Then we will discuss whether the unintended negative consequences of the mainstream approaches are observable in the presented story. Finally we will quickly touch the issue of alternatives approaches to the mainstream ones.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 054 Research on Evaluation**

**O 163 - How Gatekeepers Sometimes Enhance the Effects of Evaluation Machineries in Research**

*P. Dahler-Larsen<sup>1</sup>*

<sup>1</sup> *University of Copenhagen, Political Science, Copenhagen, Denmark*

The influence of evaluation machineries (such as bibliometric indicators) upon researcher practices is a much debated issue. Many warnings have been issued followed by advice about how to curb the influence of such machineries (such as never using one indicator alone; always be specific about which indicators are used for which purposes; and always use indicators as a support, not as a substitute for human judgment). The underlying assumption is that somehow human judgment can function as a bulwark against constitutive effects of evaluation machineries.

This paper looks at situations where evaluation machineries have more influence than what would otherwise be expected and where some form of human judgment in fact enhance or multiply the effects of evaluation machineries.

Using vignettes (small case narratives) related to the Danish Bibliometric Research Indicator (in Danish: BFI), this paper looks at how gate-keepers, bridgebuilders, interlocutors, educators, committee members and others interpret the BFI and put it to use under specific circumstances in ways which enhance rather than lessen the effects of the BFI. It is argued that a key ingredient in these situations is anticipation – and co-construction – of a not-yet-constructed reality. Gatekeepers who “know the future” play a key role, of course in combination with a range of situational factors.

By developing grounded hypothesis about how and why this multiplication of effects happens under some circumstances, this paper contributes to an understanding of the interaction between “human” and “machine-like” forms of evaluation, and to an understanding of the mechanisms which undergird constitutive effects of evaluation systems in research.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 054 Research on Evaluation****O 164 - Empirically Exploring if, where, and how Undergraduate Students are Exposed to Evaluation: Implications for Creating Global Pipelines into the Profession***J. Lavelle<sup>1</sup>, N. Sabarre<sup>2</sup>, H. Umans<sup>2</sup>*<sup>1</sup> *University of Minnesota-Twin Cities, Organizational Leadership- Policy- & Development: Evaluation Studies, Minneapolis, USA*<sup>2</sup> *Claremont Graduate University, Social Science- Policy- & Evaluation, Claremont, USA*

Scriven (2003) predicted radical changes in higher education systems away from stereotypical social science education in favor of the logical and practical application of evaluation. More than a decade later, the formal education of evaluators typically occurs at the graduate level (LaVelle, 2014; LaVelle & Donaldson, 2010), and it appears that few undergraduate courses focus on evaluation (LaVelle, Sabarre, & Uhman, in preparation), though previous research has suggested that undergraduate students are a potential pool of future evaluators (LaVelle, 2011). One of the espoused goals of the undergraduate curriculum is an informed, democratic populace (Rhodes, 2001), which has generally drawn from a liberal arts education pedagogy, and included a wide range of topics such as biological and natural sciences, social sciences, mathematics, philosophy, literature, and language arts. Some might argue, however, that the implicit goals of the undergraduate experience might be shifting away from critical thinking alone towards preparing graduates for action (Rhodes, 2001). Action, for example, in the workplace, the community, the nation, and in the global society. The merits of such a shift towards action can be debated, though it raises the question of where, if at all, evaluation currently sits in the undergraduate curriculum.

The authors will share the processes and preliminary results of an ongoing study that systematically analyzes evaluation's place in the undergraduate curriculum. The study is grounded in online curricular exploration and analysis methodologies established in previous studies (e.g., LaVelle, 2014; LaVelle & Donaldson, 2010, 2015), and is focused on the undergraduate curriculum at the top 100 colleges and universities in the United States as identified by the US News and World Report. US News and World Report has been selected as the established sample frame because its inclusion criteria are both publicly available and defensible. The resulting lists of Top 40 public and private Colleges and Universities have painted a consistent picture of excellence in the United States. The authors will describe the inductive approach to analyzing the undergraduate curriculum and evaluation's current place in preparing the next generation of democratic citizens for local/global action. Implications for the evaluation profession, practitioners, and evaluation education programs will be discussed, with a specific focus on the implications for educational systems in Europe.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 055 Values and Accountability****O 165 - Privilege Cognizant Identity and Evaluation***J. Hall<sup>1</sup>*<sup>1</sup> *University of Georgia, Department of Lifelong Education- Administration- and Policy, Athens, USA*

Broadly speaking, the field of evaluation supports the inclusion of marginalized groups in evaluation practice to address social, economic and political inequalities. Evaluations in the U.S. that aim to address inequalities include a family of stances referred to as social justice-oriented approaches (Mertens & Wilson, 2012). These approaches include, but are not limited to responsive (Stake, 1975), culturally responsive (Hood, Hopson & Frierson, 2015), values-engaged (Hall, Ahn, & Greene, 2012), participatory (Chouinard, 2013), democratic (House & Howe, 1999), empowerment (Fetterman, 2013), transformative (Mertens & Wilson, 2012) and indigenous frameworks (LaFrance & Nichols, 2010). Overall, these approaches prioritize stakeholders' views and participation in evaluation practice (Heron & Reason, 1997).

However, social-justice evaluators need to be attentive to the ways in which the notion of social justice itself shifts or is even relevant when used in various contexts. For instance, in the U.S., social justice typically aligns with democratic values, yet these values may be inappropriate in other contexts. Further, social-justice approaches are likely to garner superficial understandings of inequalities if they solely focus on marginalized groups. This is because the attention is on one side of inequality: oppression. As Pease (2010) noted, oppression reflects the "receiving end" of inequality. Overemphasizing oppression overlooks the other side of inequality: privilege. As a result, the privileged group is normalized and not scrutinized.

Therefore, I argue more engagement with privilege is needed to enhance evaluation practice. Specifically, this paper invites evaluators to consider their privilege as members of a professional group (Bailey, 1998; Pease, 2010). This is because being a member of a professional group often grants more privilege than one's personal competencies (Pease, 2010). For instance, evaluators are often assumed to have methodological expertise and, as a result, may not be questioned about their capacity to conduct an evaluation. With these points in mind, I offer the notion of "privilege cognizant identity" as one way for evaluators to critically reflect on their privilege. Privilege cognizant identity posits that members of a dominant [professional] group can: 1) learn from those in marginalized positions, 2) criticize their own privilege, and 3) act in ways that depart from their privileged social position (Bailey, 1998). Privilege cognizant identity is important to the field of evaluation because evaluators benefit from their social location and play a role in the production of privilege – whether done consciously or unconsciously. Consequently, evaluators have a societal obligation to be acutely aware of the oppressive potential of privilege. At the same time, this work recognizes the intersectionality of privilege, which refers to individuals occupying varying degrees of oppression and privilege within a socially-constructed and inequitable system (Collins, 1991). Thus, the objectives of the paper include: briefly exploring evaluation as a profession and evaluator professionalism (Duggan & Bush, 2014; Schwandt, 2015); linking evaluator professionalism to privilege (Kirkhart, 2010; Kirkhart, 2015; Pease, 2010); presenting background on and a current characterization of privilege cognizant identity, including a discussion on the intersectionality of privilege; and offering ways social-justice oriented evaluators can foster their own privilege cognizant identity.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 055 Values and Accountability

## O 166 - Closing Civic Space: Formative Research to Strengthen Grant-making Practices and Grantee Resilience

*A. Miranda*<sup>1</sup>

<sup>1</sup> *Transparency and Accountability Initiative, Learning, Washington, USA*

A vibrant, legitimate, and well-informed independent civic sector is critical to achieving development objectives across all sectors. Yet civil society groups worldwide increasingly face legal and political constraints on their ability to operate freely and independently. Since 2012, more than 100 laws have been proposed or enacted by governments that restrict the registration, operation, and funding of NGOs.

The Transparency and Accountability Initiative (TAI) is a collaborative comprising leading funders of transparency, accountability, and participation efforts worldwide. Meaningful civic participation is particularly crucial to the sustainability and viability of the transparency, accountability, and participation (TAP) agenda: to effectively hold power holders to account, non-governmental organizations (NGOs) must have a legitimate and recognized voice in society.

Threats to civic space have attracted much philanthropic and donor agency discussion. And there are active global research initiatives to monitor civic space conditions. Yet these initiatives have not yet explored the experience and needs of groups explicitly pursuing TAP outcomes. To address this gap, TAI commissioned formative mixed-methods research to better understand how members' grantees pursuing TAP outcomes are experiencing and responding to shrinking civic space.

The research objectives are to inform TAI donor member efforts to (1) strengthen TAP grantee efforts to mitigate urgent threats to civic space, and (2) support TAP field adaptation and resilience in the face of longer-term vulnerabilities. We fielded a survey with TAI member grantees, using a census sampling approach. In the coming months, we will draw case selection criteria inductively from the survey findings, and then conduct remote and field-based grantee qualitative data collection to conclude this phase of the research.

Using survey findings, the presenter will address the following questions among others: to what extent are grantees facing closing civic space; how does that experience differ by grantee goals (i.e., do TAP groups pursuing gender or equity and diversity goals experience closing civic space differently than groups pursuing other aims); how do grantees report closing civic space affecting their TAP work? Drawing on qualitative data, the presenter will further explore grantee responses to changes in civic space, and explore the implications of changes in civic space for TAP grantee/funder relationships. The presenter will also discuss learnings from the evaluation experience and any emergent next steps inspired by the findings.

The topics of closing civic space and NGO resilience are of general public interest, and the extent to which these phenomena present similar or different challenges across development sectors is an ongoing discussion to which this research may contribute. This presentation will cover themes and capabilities relevant to the evaluation community, including research of sensitive topics, navigation of power dynamics between the evaluation commissioners and the evaluand, and engagement of research stakeholders in the interpretation and use of findings.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 055 Values and Accountability**

**O 167 - System-Wide Accountability for Gender Equality  
in the United Nations System**

*P. Alvarez*<sup>1</sup>

<sup>1</sup> UN Women, Independent Evaluation Office, New York, USA

Systems thinking combined with an accountability framework to measure institutional performance in the United Nations system has allowed for an interesting monitoring exercise that has culminated with a self evaluation process and an upcoming formal evaluation.

As part of a system-wide effort to increase accountability on gender equality, in early 2012, the United Nations agreed on the landmark UN System-wide Action Plan on Gender Equality and the Empowerment of Women, or UN-SWAP, to implement the gender equality policy of its highest executive body, the UN Chief Executives Board, chaired by the UN Secretary-General.

Spearheaded by UN Women, the UN-SWAP for the first time assigns common performance standards for the gender-related work of all UN entities, ensuring greater coherence and accountability. After 5 years of implementation, a new framework has been developed building on lessons learned and an evaluative assessment of correlations and drivers for success. The UN-SWAP 2.0 extends the reach of UN-SWAP 1.0 by including new Performance Indicators on gender-related and SDG relevant results, as well as a new one on leadership.

The presentation will focus on the systemic aspects of this framework and the evaluative analysis of its strengths and limitations as an institutional performance framework covering up to 66 diverse organizations within the UN system. Mix-methods, the indicator technology reveal the underlying values of the international bureaucracy and the complexity of its institutional structures.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 056 Evaluating Public Health Initiatives

## O 168 - Evaluation of the Inspection and Surveillance Model for Access, Use and Quality in the Provision of Health Services in Colombia

*R.E. Penaloza Quintero<sup>1</sup>, L. Mina Rosero<sup>1</sup>, V. Suelt Cock<sup>1</sup>*

<sup>1</sup> Pontificia Universidad Javeriana, Distrito Capital, Bogota, Colombia

The objective of the evaluation was to know if the Surveillance and Control Inspection model developed by the Ministry of Health of Colombia and the National Health Superintendence, responded to the main operating problems of the Colombian health system.

The evaluation focused on reconstructing the theory of action of inspection and surveillance against the main problems of operation of the system. At the same time, reconstruct the causal theory of the main problems based on the perceptions of the different actors in the system.

The methodological procedure was developed in three stages: in the first stage, a matrix was created with functional categories of health system (Financing, Insurance, Provision of services, among others) and subsequently the policies, norms and other instruments used by the government to perform the inspection and surveillance. Faced with this matrix, the problems that it sought to solve with these actions were defined. (Defining the problems that the intervention model addressed)

In the second stage, it was identified, through interviews with key actors and workshops developed with different institutions, a set of existing problems that should be inspected and monitored, and a matrix with these problems was built.

Finally, a comparison is made of the problems identified by the actors and the problems that were addressed by the inspection and surveillance model and the relevance (external coherence) of the model was established in relation to the perception of the actors.

**Results:** There is no relevance among the main problems identified by the different actors in relation to the problems that the inspection and surveillance model of the Ministry of Health and the National Health Superintendence addresses. From the perception of the actors, the main problems were: Quality service network organization, Health risk management, Affiliation, Representation of affiliates and Financial risk management.

Faced with these problems, there are no inspection and surveillance mechanisms on the part of the inspection and surveillance model.

In the same way, it is evident that the activities of the model respond to problematic situations and not to control the causes that generated them. In the identification of the problems it was evidenced that it was necessary to move from a reactive to a preventive inspection model. Develop inspection and surveillance actions against possible risks and not in front of problems.

The result of the evaluation made it possible to define new actions that should be developed in the inspection and surveillance model to avoid problems in the operation of the health system.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 056 Evaluating Public Health Initiatives****O 169 - Application of the Constructivist Evaluation to Behavioral Health Issues: Facilitating the Placing of Value by Others**G. Tchamba<sup>1</sup><sup>1</sup> ADR Evaluation Consulting, NA, Berrien Springs, United States Minor Outlying Island

The valuing branch of the evaluation theory tree assigns the role of placing value to evaluators (Alkin and Christie, 2004). However, some theorists of the valuing branch, specifically the Fourth Generation Evaluation (FGE) theorists, view the placing of value as a shared task between evaluators and stakeholders (Guba and Lincoln, 1989). The FGE or constructivist evaluation argument is based on constructivist philosophy which argues that reality is socially constructed. Critiques of this evaluation approach argue that there are discrepancies between this approach and the field-based experience involving evaluations done under contract (Lai, 1991). As a result, evaluators seldom use the Fourth Generation Evaluation. This lack in the evaluation literature has further widened the gap between this theory and the frequently used approaches such as participatory, developmental, empowerment, etc. However, evaluators who are interested in understanding multiple perspectives on program outcomes will benefit from the use of the constructivist evaluation.

In order to address the utility of the constructivist evaluation, the sole purpose of this paper is to present the application of this approach to the understanding of primary care providers' perspectives on the health benefits of moderate alcohol consumption. This is based on the evidence in epidemiological literature that moderate drinkers (14–21 drinks a week) have better health and longer life expectancy than those who abstain (Nova et al, 2012).

The audience at this presentation will learn concrete ways to navigate through the two main phases of this approach (discovery and assimilation). The findings are discussed and emphasis is placed on the variable of "doubt" identified from the analysis of the interviews with primary care providers and illustrated by the "theory of conflict", (a theory built upon the experiences of primary care providers that try to make sense of the relationship between moderate alcohol consumption and health). Finally, the paper discusses the implication for evaluators in behavioral health while suggesting the scope for further evaluation with larger sample of primary care providers.

This paper reveals how real life application of this approach is significant and useful for the following reasons: 1) it adds to the qualitative literature, specifically the constructivist grounded theory which defines well the idea of "facilitating the placing of value by others" as the theorists intended. 2) It can serve as guide to evaluators in behavioral health seeking to employ constructivist methodology in their programs. 3) It can ignite the contemporary discourse on the use of constructivism in evaluation. 4) The knowledge gained from this application has value as it augments the increasing number of evaluators interested in meaning making in evaluation practice.

**Reference:**

Alkin, M. C., & Christie, C. A. (2004). An evaluation theory tree revisited. In M. C. Alkin (Ed.), *Evaluation roots*. Thousand Oaks: Sage.

Guba, E., & Lincoln, Y. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage Publications. Lai, Morris K. (1991). *Field-based Concerns About Fourth-Generation Evaluation Theory*

Nova, E., Baccan, G., Veses, A., Zapatera, B., Marcos, A. (2012). Potential health benefits of moderate alcohol consumption: current perspectives in research. *The Proceedings of the Nutrition Society*; 71(2):307–315

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 062 Improving Systems and Capacities through Evaluation****O 170 - A Novel Methodology to Examine the Know-Do Gap in Childhood Pneumonia Assessment Among Frontline Health Workers in an Indian State***L. Ray Saraswati<sup>1</sup>, A. Mishra<sup>1</sup>*<sup>1</sup> RTI International - India, Public Health, New Delhi, India

**Introduction:** In India, frontline workers (FLWs) – accredited social health activists (ASHAs) and rural medical providers (RMPs) – being the first point of contact of caregivers, play a pivotal role in early detection and prompt treatment of childhood pneumonia. Present study adopts a novel methodology to assess the gap between current knowledge and assessment skills related to management of childhood pneumonia among ASHAs and RMPs, and identify intervention strategies that could be used to reduce the gap.

**Methodology:** As part of a study to inform an intervention focused on reducing under-five mortality, we surveyed 473 ASHAs and 447 RMPs in six districts of Uttar Pradesh. While their knowledge was assessed using face-to-face interviews, their assessment skills were assessed using video vignettes. 'Shortness of breath' and 'chest in-drawing' were considered key signs of pneumonia. We used binary logistic regression (separately for ASHAs and RMPs) twice, first to assess the effectiveness of different information, education, and communication strategies in improving the knowledge of FLWs; and second, to identify strategies that are effective in improving the assessment skills of a subset of FLWs that knew about the signs.

**Findings:** Although 50.7% ASHAs and 41.6% RMPs reported to know at least one sign of pneumonia, 12.8% ASHAs and 11.3% RMPs could identify the signs from video vignette. ASHAs showed significant improvement in knowledge if they had met with a community resource person [Odds Ratio (OR)=1.97; 95%CI=1.00–3.91], or had received training on ASHA module [OR=2.58; 95%CI=1.54–4.34], or had obtained pneumonia-related information from television [OR=1.67; 95%CI=1.11–2.53]. We could not identify any factor that significantly affected the assessment skills of ASHAs. Among RMPs, pneumonia related inter-personal communication from a non-government health worker [OR=2.14; 95%CI=1.31–3.52], pneumonia related information from television [OR=2.28; 95%CI=1.45–3.60], and personally seeing at least one child with pneumonia in the last week [OR=1.64; 95%CI=1.00–2.70] significantly contributed to their knowledge regarding the signs. Among those RMPs that knew about the signs, being in older age-group significantly improved the assessment skills.

**Conclusion:** Though FLWs have reasonable knowledge about pneumonia signs, they lacked on assessment skills required for appropriate case management of childhood pneumonia, resulting in a know-do gap. While some of the existing intervention strategies seemed to work in increasing FLWs' knowledge, none was effective in improving their assessment skills. Programmes should explore innovative strategies for improving the assessment skills of FLWs, besides focusing on the existing knowledge enhancement strategies like frequent meetings with community resource person and training on ASHA module for ASHAs, messaging through inter-personal communication for RMPs, and use of mass-media for both.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 056 Evaluating Public Health Initiatives

## O 171 - Very Low Birth Weight Newborn Tracking Assessment: An Evaluation Evolving into Ownership Transfer

A. Das<sup>1</sup>, P. Maharana<sup>1</sup>, T. Mahapatra<sup>1</sup>, S. Srikantiah<sup>1</sup>

<sup>1</sup> CARE India Solutions for Sustainable Development, Concurrent Measurement and Learning, Patna, India

**Rationale:** In absence of well-equipped neonatal care facilities serving rural Bihar, practice of essential home-based newborn care assumes supreme importance towards reduction of neonatal and infant mortality, especially that happening among pre-term and small-for-gestational age (Low birth weight) newborns. Therefore, it was hypothesized that scaling up the coverage of essential services to high-risk newborns at community-level could prevent adverse outcomes among them. With the long-term objective of improving the health parameters among Bihar's children, Bihar Technical Support Program (BTSP), led by CARE India and financially assisted by the Bill & Melinda Gates Foundation, launched the "Very Low Birth Weight (vLBW) Tracking" intervention in 2015.

**Objectives:** To assess the effectiveness of the intervention in improving the coverage of advices on essential newborn care and for bringing down the morbidity and mortality, 200 public health facilities were randomly selected from the list of all functional delivery points in public sector in Bihar and various intermediate and distal outcomes were tracked among the neonates born in those facilities. Thus far, four rounds of assessment have been conducted – one during the pre-intervention period (neonates born during Feb-Apr'2015) and three during post-intervention period (for children born during Sep-Oct'2015, Sep-Oct'2016 and Sep-Oct'2017).

**Findings:** Significant improvements between baseline and post-intervention rounds were noted regarding the % of mothers who were informed that their children were 'Weak' – at the facility (R1: 25%, R2: 54%, R3: 60% and R4: 50%) as well as during home visit by FLW (R1: 19%, R2: 48%, R3: 43% and R4: 30%). Practice of "Kangaroo Mother Care (KMC)" – an important component of essential newborn care – showed significant improvement in post-intervention period compared to baseline in both facility (R1: 15%, R2: 14%, R3: 39% and R4: 31%) and home (R1: 10%, R2: 15%, R3: 37% and R4: 29%). An increasing trend was noted regarding detection and birth weight recording of the extremely low birth weight newborns (<1500 g) showed an increasing trend – which, on its own, can be considered a vital success of the intervention. Moreover, a downward trend in mortality across the rounds of assessments, in each strata of birth weight (<1500 g, 1500 – 1799 g and ≥1800 g) can be noted. Therefore, to adjust for the differential distribution of birth weights in different rounds, it was decided to statistically adjust for the influence of birth weight. Following such adjustment, the birth weight adjusted mortality was found to decline significantly from the R1 (22.11%) to R4 (11.87%). The significantly declining trend was also observed for both early and late neonatal mortality, as well as for morbidities during the neonatal period. Multiple regression analysis was undertaken to determine the predictors of survival during neonatal period among vLBW newborns and following determinants were found to be associated with better survival – birth during the immediate post-intervention phase and that during the maintenance phase, birth weight higher than 1500 g, children of low parity mothers, receiving visit from FLW in the first week and/or receiving advice on extra care from FLW.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 056 Evaluating Public Health Initiatives

## O 172 - Zika Outbreak in Brazil: Understanding Perceptions to Increase the Effectiveness of Preventive Action

*C. Santos<sup>1</sup>, V. Mansour<sup>2</sup>*

<sup>1</sup> *Plan Eval, Research, São Paulo, Brazil*

<sup>2</sup> *Food and Agriculture Organization of the United Nations FAO, Office of Evaluation, Rome, Italy*

Transmitted by the *Aedes aegypti* mosquito, the Zika virus has become a cause for global concern due to its connection with neurological abnormalities in newborn children, especially microcephaly (congenital Zika syndrome). Brazil's outbreak began in 2014, and by 2015 there were anywhere between 440 thousand and 1.3 million cases of Zika, and nearly 5 thousand suspected cases of microcephaly.

Preventive measures to reduce the number of mosquitoes in Brazil first emerged in the fight against Dengue in the 1980s. Most of these strategies focused on the insect instead of fostering individual behavior change (e.g. removing stagnant water) and did not succeed in eradicating the disease.

When it comes to a potentially life-threatening virus like Zika, community-level action is crucial to reducing the risk of infection. In comparison with approaches that entail a large-scale overhaul of physical infrastructure, preventive measures can promote faster and more effective responses, and communication strategies play a central role in promoting transformative changes among individuals and groups.

This paper reports the findings of research conducted in the cities of Campina Grande and Recife in March 2016 in order to identify optimal social communication strategies (e.g. UNICEF's "Communication for Development") for increasing awareness around Zika, which at the time was still infecting many people. Following an extensive desk review of past research and information campaigns on other mosquito-borne diseases, we held focus groups with key public officials (community health workers and epidemiologists), residents of affected communities, and members of the most vulnerable demographic groups (i.e. teens and women of childbearing age).

Despite the large number of information campaigns about mosquito-borne diseases, the failure to emphasize preventive measures meant that the public remained largely uninformed about mechanisms of transmission and telltale symptoms of infection. For instance, people doubted that mosquitoes could transmit Zika, and did not see how stagnant water could be ideal for mosquito reproduction. Furthermore, individuals felt they had little to gain from taking preventive measures, and rarely did so.

However, the most vulnerable members of society tended to be more proactive. Given the serious consequences for unborn children, women led the way in changing behavior at the individual and community levels. Participants also believed teens play an important role in spreading awareness.

This research provides local governments and the international community with insights into how information campaigns can promote transformative changes in response to an emergency. By investigating public perceptions of the transmission and risk associated with Zika, our study informed the creation of communication tools that encourage individuals to incorporate preventive measures into their daily routines.

As viruses like Zika mutate and evolve, countries will have to rely on effective and efficient public health strategies that help citizens stay informed, and build resilience at a societal level. Therefore, we consider this study to be thematically and methodologically relevant for the international evaluation community.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 057 Institutionalization of Evaluation Policies**

**O 174 - Development of National Evaluation Systems in Africa –  
A Comparative Study of Ghana and South Africa**

*C. Amoatey<sup>1</sup>, C. Blaser<sup>2</sup>*

<sup>1</sup> Ghana Institute of Management and Public Administration GIMPA, Management Science, Accra, Ghana

<sup>2</sup> Center for Learning on Evaluation and Results in Anglophone Africa CLEAR AA, University of Witwatersrand, Johannesburg, South Africa

Ghana and South Africa are two countries pioneering an agenda on evidence use in the region. South Africa's DPME developed a national evaluation policy framework, which was rolled out from 2012 (Goldman, et al, 2015). Ghana is in the process of developing its own national evaluation policy. However, this policy development is happening in a different historical and democratic context. A better understanding of how these monitoring and evaluation systems have emerged in South Africa and Ghana will help us better understand the enablers and constraints to using evidence in both contexts. The study show that evaluation systems of countries are strongly influenced by their political economies. At the same time, CLEAR-AA is developing a working definition of monitoring and evaluation systems that will help systematize our research on where the unevenness in monitoring and evaluation systems lie in the region, and how they interact. Bringing these two areas of enquiry together will help inform how best to strengthen evaluation systems capacity in the region.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 057 Institutionalization of Evaluation Policies

## O 175 - Moving from a Decentralised to a Centralised Corporate Evaluation System: Opportunities for Methodological Innovation Combining Contribution Analysis and Oecd/Dac Criteria

*V. Hundt<sup>1</sup>, L. Bendfeldt-Huthmann<sup>1</sup>*

<sup>1</sup> GIZ - Deutsche Gesellschaft für Internationale Zusammenarbeit, Evaluation Unit, Bonn, Germany

In order to enhance substantially the quality of project implementation, evaluations need to assess not only results but also the causal relationships between a projects activities, outputs, outcomes and impacts. The results of such analyses enhance evidence-based recommendations and learning processes. Therefore, the evaluation unit of GIZ – the German state-owned implementing agency for international cooperation – has put focus on improving the analysis of contributions of the projects within the framework of the OECD/DAC criteria.

In order to monitor and improve evaluation quality, GIZ carried out meta-evaluations of the project evaluations in 2015 and 2016. The meta-evaluations showed that particularly the methodological quality of the project evaluations should be improved while process quality and utility were already quite high. By then, the project evaluations within GIZ were implemented on a decentralised basis. The corporate evaluation unit set the standards but the project leader were responsible for contracting the evaluators and organizing the process. Also, the evaluations were combined with the appraisal mission of a follow-on phase. Although this combination of both processes proved useful for knowledge transfer, it resulted in too little concentration on each of the both processes, particularly regarding methodological quality.

In 2018, a new Evaluation Policy of GIZ was published with the stipulation of a new evaluation system. The former decentral project evaluations are replaced by central project evaluations (CPE) steered by the corporate evaluation unit. During the pilot phase starting in October 2017, six evaluations were already carried out based on a random sample. In 2018, about 40 central project evaluations are planned. In 2019, the system will be fully running with about 100 central project evaluations per year, based on a random sample.

One specific goal of the new evaluation policy is the improvement of evidence regarding specific contributions of projects' activities and outputs to outcomes and impacts. Therefore, as a minimum standard, a contribution analysis is introduced as part of the assessment of the OECD/DAC criteria. Based on the projects' results model, external evaluators select the main results hypotheses of the project. The selected results hypotheses are analysed and assessed as specific evaluation dimensions under the OECD/DAC criteria effectiveness and impact.

Experience from the pilot evaluations shows that while this approach is very useful it also leads to a number of challenges. Therefore, we would also like to discuss current challenges and approaches regarding: (1) Selection of results hypotheses and its influence on overall assessment of effectiveness and impact, (2) Handling of alternative hypotheses given only one evaluation mission, (3) Simple presentation of results of the contribution analysis to a variety of stakeholders.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 058 Defining Objectives and Indicators**

**O 176 - A Methodological Framework for the Evaluation and Selection of Policies for Endogenous Regional Development. The Case of Greek Regions**

*P. Koudoumakis<sup>1</sup>, G. Botzoris<sup>1</sup>, A. Protopapas<sup>1</sup>*

<sup>1</sup> *Democritus University of Thrace, Civil Engineering, Xanthi, Greece*

Regional development has been extensively studied over the past decades. Research is often characterized by emphasis not only on the type of the development but also, on the selected policies that need to be implemented so that endogenous development can be achieved. Endogenous is a form of regional development, where local factors, such as human capital, enterprises, and local initiative, are the main drivers of the developmental process. The fundamental reason for the contentious debate among scientists – regarding both the results of the regional development policies and observed inequalities in the level of development – is the absence of a commonly accepted methodological framework for evaluating and selecting regional development policies.

In this paper, an attempt is made to develop a new methodological framework in order to construct composite indicators, for the evaluation and selection of policies that will drive endogenous regional development. Composite indicators are increasingly used due to their capability to integrate many individual variables, by providing a synthesis of multidimensional issues that can be easily understood by regional development policymakers. However, review of the relevant literature has revealed that composite indicators are often based on empirical approaches and specific case studies and therefore their applicability is limited to the specific problem for which they have been developed.

The proposed methodology, has been developed by a combination of techniques designed to solve research challenges which involve (a) the estimation of missing data in a time series of variables, (b) the evaluation of the statistical significance of variables associated with an indicator, c) the aggregation of variables, a challenge that coexists with the issue of weighting and d) normalization of data.

Based on widely accepted theories of regional development and a significant data inventory for the thirteen Greek Regions, twenty-two variables in five key areas have been chosen. More specifically, the analysis includes fields which are considered to be developmental factors of a region's productive system like 1) human capital, 2) research and development, 3) industry, trade, transport, and 5) tourism and the service sector.

The proposed method of constructing composite indicators includes the implementation of the Structural Equation Models (SEM) approach to estimate the statistical significance of the variables which compose the indicator. Subsequently, the values of the composite indices which have been constructed on the basis of the Principal Component Analysis (PCA), were used as input to the Data Envelopment Analysis (DEA): (a) for the assessment of regional developmental policies and (b) as a tool for selecting regional developmental policies by identifying the necessary improvements in productive resources so that inefficient regions do manage to become efficient. The analysis demonstrates that, with better use of their most productive resources, the Greek Regions may increase their average GDP by approximately around 8%. An area for developmental interventions by policymakers might be the optimized utilization of human resources since they lead to a significant improvement in efficiency.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 058 Defining Objectives and Indicators**

**O 178 - Evaluating the United Nations' Sustainable Development Goals: Dropout Prevention Program in Finnish Vocational Education**

V. Vehkasalo<sup>1</sup>

<sup>1</sup> National Audit Office of Finland, Performance audit, Helsinki, Finland

The United Nations' 17 Sustainable Development Goals (SDGs), adopted in 2015, aim to enhance wellbeing and resilience globally by 2030. The goals are quite general, for instance “end poverty” and “end hunger”. The official SDG indicators which aim to monitor the progress towards these goals are usually expressed as a proportion of population under certain threshold levels – for instance, the proportion of population below the international poverty level. However, measuring just the progress of these indicators is somewhat misleading, if we wish to make evaluative statements on the success of specifically designed programs aiming towards these goals. As always, there may be several confounding factors which also have an effect on the goals, besides the program(s) under study. Therefore in order to find conclusive evidence of program effectiveness, we have to utilize more elaborate methods.

As a case study we present a recent (2018) performance audit, which evaluated the effects of dropout prevention program in Finnish vocational education. Although this particular program preceded the SDGs, it was directly related to SDG 4, quality education, and target 4.4, which aims to increase the number of youths who have relevant vocational skills for employment. Dropping out has been a major problem in Finnish vocational institutions, and in 2011–2014, a large-scale dropout prevention program was carried out in order to increase completion rates. We evaluated the effects of the program using register data on students (N = 24,000) and difference-in-differences estimators. According to our estimates, the program did not have any effect on completion or dropout rates. Dropout rates decreased and completion rates increased, but these changes were not caused by the program. In fact, these changes were larger among the control group students. Likely reasons include Finland's prolonged economic recession and tightened criteria for youth unemployment benefits.

Note that the program in question started a few years before the SDGs were officially adopted. Nevertheless, it is a good example of the kind of programs that we might also later use in order to reach the SDGs. The lesson to learn is that if we only had relied on indicator-type monitoring – for instance, the dropout rate of vocational students – our evaluation on the effects of the program would have been erroneous. We would have concluded that the program decreased dropouts, and hence benefited towards the SDGs, while the opposite is true. Therefore we must constantly bear in mind what we are measuring and how.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 059 Evaluation Planning and Evaluability Assessment**

**O 179 - Repurposing the Qsort Method and the Classic Simulation Method to Isolate Criteria That Underlie Programme Evaluability Decisions**

A. Boodhoo<sup>1</sup>

<sup>1</sup> *University of Cape Town, Management Studies, Cape Town, South Africa*

Evaluators work within messy and dynamic environments, characterised by real-world complexities and constraints. They continuously engage in a “complicated juggling act involving trade-offs between available resources and acceptable standards of evaluation practice” (Bamberger, Rugh, & Mabry, 2012, p.7). Our practice rarely complies with prescriptive theories of how evaluations ought to be conducted. Instead, it typically calls for a less formulaic approach, whereby practice decisions are negotiated in a discretionary space and driven by ‘implicit’ theories that account for the presenting features of evaluation contexts. It is imperative that we investigate systematically what ‘implicit’ theories exist around different areas of practice to consolidate our understanding of what actually unfolds on the ‘rough ground’.

In this study we focus on an under-investigated but fundamental segment of our practice: deciding on the evaluability of a programme. We use an adapted version of the conventionally onerous Q sort method to explore inductively how a large and geographically dispersed sample of evaluators operationalize programme evaluability. While this method had been used in programme evaluation before, its application is rare as the limits and mechanics of this method are not well understood. We also use a scenario-based task to assess whether or not evaluators’ operationalizations are consistent across evaluation contexts. By manipulating systematically selected evaluability conditions within fictitious evaluation scenarios, it is possible to examine how and if evaluators reshape their operationalizations of programme evaluability depending on the features of the evaluation context. We also creatively combine the results derived from each method to address the following question: Do evaluators use the same criteria prioritised in the Q Sort task to guide their programme evaluability decisions, when confronted with scenarios that mimic real evaluation situations.

We discuss the principles underlying the design and administration of each method and engage in a discussion of how these methods could be further refined and supplemented for wider application in the study of evaluation practice.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 059 Evaluation Planning and Evaluability Assessment**

**O 180 - Design evaluation of the Master Plan for International Cooperation for Development**

*S. Ulla Díez<sup>1</sup>, R. Poo<sup>2</sup>, P. Caballero<sup>3</sup>*

<sup>1</sup> Ministry of Foreign Affairs and Cooperation of Spain, General Directorate of Development Sustainable Policies, Madrid, Spain

<sup>2</sup> Independent, Madrid, Spain

<sup>3</sup> Ministry of Foreign Affairs and Cooperation of Spain, General Directorate for Sustainable Development Policies, Madrid, Spain

The Master Plan for Spanish Cooperation responds to Law 23/1998 on International Cooperation for Development which establishes that its function is to establish and plan the priorities of the policy for this area in each period.

Based on the known and reported difficulties to evaluate these plans, and their limitations to serve as effective guidelines for action planning, an evaluation of the design of the Master Plan was carried out by the Division of Evaluation and knowledge Management of the Ministry of Foreign Affairs and Cooperation of Spain. The primary purpose of this paper is to illustrate lessons learned about design evaluation and its added value for policy design.

Policy design and evaluability are often highlighted in academic literature, but in the reality of policy design, this is not always taken into consideration. And there are not many design evaluations to promote changes in that direction.

In this evaluation, the aim was to identify the strengths and limitations in the formulation of the IV Master Plan as a strategic planning document, and to provide useful information for the articulation of the V Master Plan, which had to be prepared. An evaluability assessment was designed with a specific focus on design aspects. The dimensions analyzed were: Rationality, internal coherence, identification of assumptions and risks and empirical support to the objectives. Conclusions and recommendations will be presented with a clear focus on lessons learned applicable to other policies as well as other design evaluations.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 059 Evaluation Planning and Evaluability Assessment**

**O 181 - Use of Conversational Mapping for Evaluative Processes**

*E. Canu<sup>1</sup>, S. Pistis<sup>1</sup>*

<sup>1</sup> Regione Autonoma della Sardegna, Nucleo di Valutazione e Verifica degli Investimenti Pubblici, Cagliari, Italy

Aim of this contribution is to discuss how the tool “conversation mapping” can be used in a participatory process to elicit stakeholders’ views on specific topics.

“Conversation mapping” is a participatory tool that enable its users to track and eventually model multiple perspectives held by different stakeholders about a given topic /problem.

The process consists in holding a conversation (between 4–9 participants) about a salient aspect of the problem (the “trigger”) while writing on a “map” the essence of their contributions. This allows to compare and contrast different positions and understandings of the situation, and gaining insights.

The ultimate goal is to share a new, more complex perspective of the problem, and to generate and share common positions.

Two experience will be presented that will highlight both strenghts and drawbacks of this technique: the first, the 2013 meeting “Measuring wellbeing in Newcastle” (part of the OECD project “Measuring Well-being and Progress”); the second one, a technical meeting held by the Italian Evaluation Units Network in Cagliari in 2018, concerning the model of governance of Evaluation Plans and the management of the tendering process.

Some caveats will be provided as well, specifically the need to maximise diversity to enable a diversity of views to be expressed.

It will then be discussed its applicability to evaluation in different stages of the evaluative processes (mainly planning and formulation of evaluative questions).

The conclusion is that conversation mapping does not produce “ready to use” evaluative material, but can help evaluators to shape better, more focused evaluative questions and to strenghten the sense of ownership held by the stakeholders.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 060 Evaluating Peace Processes**

**O 182 - The Changing Logic of the Northern Ireland Peace Programmes**

*T. Dignan*<sup>1</sup>

<sup>1</sup> *Economic Research and Evaluation, None, Belfast, United Kingdom*

This paper will provide a critical examination of the changing 'intervention logic' of the European Union's programmes to support the Northern Ireland peace process. The primary objective is to highlight the challenges faced in evaluating the impact of that support, which is now in its fourth iteration. The first EU programme (Peace I) was launched in 1995, in response to the ceasefires announced by the paramilitary organisations in Northern Ireland in the autumn of 1994. The strategic aim of the first Peace programme, which ran from 1995 to 1999, was to "reinforce progress towards a peaceful and stable society and to promote reconciliation". The programme was designed around a standard socio-economic development model, comprising a mix of social inclusion and economic measures. The distinctive features of Peace I derived largely from the attempt to embed a set of 'principles' in the implementation and delivery mechanisms; such as fostering local partnerships to facilitate political engagement between divided communities and targeting on those most affected by the conflict which had raged for the previous 25 years. In particular, social inclusion was cast as 'the pathway to reconciliation'. However, that first programme did not specifically define 'peace' and/or 'reconciliation'. It therefore proved difficult to evaluate the impact of the funding on the strategic peace and reconciliation aim. The second programme (Peace II) was implemented over the period 2000 to 2006. While continuing with a socio-economic thrust in the programme design, Peace II sought a firmer focus by requiring applications for funding support to demonstrate the 'distinctiveness' of their proposed project in terms of how it related to the strategic peace and reconciliation aims. Towards the end of Peace II, discussions around programme design became more strongly shaped by debates around peacebuilding and associated 'theories of change'. Those discussions were reflected in the design of the Peace III programme (2007 to 2013). In particular, Peace III placed particular emphasis on supporting actions which were intended to be more directly linked to the overall peace and reconciliation aim, including activities designed to reconcile communities through encouraging and facilitating cross-community interactions and relationships, complemented by actions designed to contribute to 'shared spaces' in the context of a society where the two main communities exhibit high degrees of segregation. Consequently, the socio-economic development thrust of the first two Peace programmes has largely been discarded. The fourth Peace programme, which commenced in 2014 and will run until 2020, has continued in the same vein. Notwithstanding their more direct focus, the evaluation of the 'community relations' model underpinning Peace III and IV remains challenging.

Regardless of the 'intervention logic', a fundamental challenge for evaluation of the Northern Ireland Peace programmes is the essentially qualitative nature of the desired outcome. The second major challenge is to take account of the wider context around the Peace programmes, especially the political environment, which poses considerable difficulties in disentangling the progress made by the Peace programmes and their contribution to the Northern Ireland peace process.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 060 Evaluating Peace Processes

## O 183 - Evaluating Peace. Flexibility and Complexity in the Evaluation of the Norwegian Engagement in the Peace Process in Colombia

A. Wilhelmsen<sup>1</sup>, J. Fabra-Mata<sup>1</sup>

<sup>1</sup> Norad, Evaluation Department, Oslo, Norway

Peace processes are complex and few have been evaluated comprehensively, especially by a stakeholder.[1] In this paper, we demonstrate that it is possible. However, it is necessary to use an approach that addresses the complexity of such processes. In the evaluation of the Norwegian engagement in the peace process between the Colombian government and the FARC (the Colombian peace process), this was addressed by adjusting the evaluation management model, employing a theory-based approach and applying a set of risk mitigation strategies.

Building peace is not straightforward. Even the best designed and planned processes are not immune to unexpected change, disorder and instability. A peace process affects the whole society and most likely entails major socio-political and legal restructuring. Peace processes, including the Colombian, pass the litmus of complexity with flying colors.[2]

Yet another layer of complexity exists: ownership and endogeneity in peace facilitation. If the parties own the process, but another country like Norway becomes part of the process, where to draw the line between the actors' actions, responsibilities and contributions? Adding to this, the evaluation took place in the implementation of the peace agreement phase, at a time with heightened debates in Colombia and during an election year in the country.

The evaluation addressed this complexity by (1) adapting the evaluation management model most frequently employed by the Evaluation Department at Norad, (2) developing a theory-based approach suited to the evaluation object and (3) adopting risk mitigation strategies sensitive to external factors.

On the former, it was decided that the evaluation would be led internally by the Evaluation Department. External experts were engaged to support specific parts of the evaluation like archival research, social media analysis and to conduct quality assurance, and securing the independence of the evaluation.

On the second, the evaluation dealt with complexity by isolating and working with selected nested system – intersecting subsystems that evolved over time – and turning points.

Lastly, several measures were taken to address potential risks. First, it was necessary to adapt the timeline to the Colombian elections. Secondly, it was challenging to reconstruct history solely based on archives, and the evaluation had to rely heavily on other primary data sources. Due to sensitivity, several access strategies and ethical safeguards were adopted in organizing and conducting interviews.

From this, we have some key takeaways: there are many ways to organize an evaluation efficiently without compromising independence, some better suited to complex sensitive evaluation object than others; internal validation is to be reinforced with multiple data types; it is key to identify risks and develop mitigation strategies early on, leaving room for flexibility along the way.

[1] At this point, we know of: Norad (2011): *Pawns of Peace. Evaluation of Norwegian peace efforts in Sri Lanka, 1997 – 2009*. Report 5/2011. Oslo

[2] See, for example, Bamberger, M., Vaessen, J., and Raimondo, E. (eds.) (2016) *Dealing With Complexity in Development Evaluation – A Practical Approach*. SAGE publications; Forss, K., Marra, M., and Schwartz, R. (eds.) (2011) *Evaluating the Complex. Attribution, Contribution and Beyond*, Transaction Publishers.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 060 Evaluating Peace Processes**

**O 184 - Transitional Justice: Views from the Ground on How Afghanistan Fares**

*C.R. Echavez*<sup>1</sup>

<sup>1</sup> *Research Institute for Mindanao Culture RIMCU, Social Science, Cagayan de Oro, Philippines*

The concept of transitional justice renews the trust between the population and the state, trust that is built as the prerequisite to both economic and social development. The study sought the views of Afghans on their preferred transitional justice policies, and whether these policies should be a precursor or linked to the peace process. Across five provinces, the study asked whether accountability for past human rights abuses remains a priority. A mixture of qualitative research methods was used, including assessing the perceptions and expectations of the study sites regarding transitional justice mechanisms and the issues of the five provinces included in the study. An extensive desk review was conducted before the collection of primary data. The data were based on extensive field research with focus group discussions (FGDs) and two rounds of in-depth interviews (IDIs) with the aid of semi-structured interview guides. The selection criteria of the five provinces were based on the following considerations: how conflict affects each province, factors such as levels of sympathy for armed groups, and confidence in the national reconciliation process. With the abovementioned criteria, the following sites were selected: Kabul, Bamyan, Nangarhar, Uruzgan, and Baghlan. The majority of the FGD participants, of both sexes and in all provinces, recognised that transitional justice is an idea related to addressing the past, although none of the participants had a detailed understanding of the term “transitional justice.” The FGD participants across all provinces expressed a clear understanding that the abuses of the past continue to affect the present and shape the future, and that impunity for past atrocities was fuelling new cycles of violence. A small number of FGD participants across all five provinces noted that the failure to address the past and provide justice for earlier atrocities had broader implications related to poor governance in the current context. The informants attributed the deteriorating security situation, ongoing violence, and continued corruption to the failure to hold individuals accountable for their illegal behaviour in the past. Across all provinces, there was an agreement among the majority of female and male FGD participants that the past should be addressed. Regarding the range of grievances, some participants recalled the human rights abuses suffered directly by them or by one of their family members. This issue was particularly common in Uruzgan, where the majority of participants revealed that they had directly experienced a human rights violation. Across all provinces, the majority of male and female participants provided anecdotal evidence of past abuses that had occurred in their area. Bamyan participants were unanimous that the Taliban’s oppression marked a very dark era. A few participants in Bamyan also mentioned well-known atrocities like the Yakawlang massacre, believing that justice must be sought for such events. In Nangarhar, several informants referred to the need to include Daesh in any transitional justice process. Informants in Baghlan also identified widespread gender-based violence against women, girls, and young boys perpetrated by the People’s Democratic Party of Afghanistan and the Mujahideen.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 061 Participatory and Stakeholder Approaches

## O 185 - Why Do Citizens not Participate in Community-based Monitoring? Insights from Uganda's Rural Water Sector

*N. Holvoet<sup>1</sup>, S. Dewachter<sup>1</sup>*

<sup>1</sup> *University of Antwerp, Institute of Development Policy, Antwerp, Belgium*

Since the turn of the century community-based monitoring (CBM) has become increasingly popular, particularly in settings where state-driven top-down M&E mechanisms are largely absent or dysfunctional. Bottom-up and citizen-led monitoring are propagated, starting from the assumption that they enhance local ownership, trigger transparency and accountability, foster local learning and contribute to improved local service delivery and natural resource management. While the popularity among practitioners has also recently led to an upsurge of academic research, there is still no unequivocal evidence on the impact of CBM. From an evaluative point of view the diverging conclusions are not entirely surprising: different studies use different (impact) evaluation methods (ranging from RCT to more qualitative studies), dependent variables vary (e.g. service delivery in sectors with different degrees of 'publicness'), local settings differ on socio-political and cultural characteristics and even if community initiatives appear highly similar at first glance, a closer look often unveils important differences, inter alia in the degree of inclusiveness of participants.

While citizen's participation in CBM initiatives is often assumed to exist automatically, it is increasingly acknowledged that this essential building block is not self-evident. However, as most impact evaluations fail to include detailed process evaluations, we often lack insight into the type and effective degree of citizen's participation and the factors that drive or restrain citizen's involvement in CBM. Drawing upon data from a Western Ugandan village where access to high quality water is a pressing community problem, this paper investigates why citizens are limitedly involved in community monitoring for better water service delivery.

Our study adopts a theory-driven approach and conceptualises CBM as collective action which generally refers to actions undertaken by a group of people with the aim to achieve a common objective. Collective action does not arise automatically and its problematic nature has led to a stream of literature from various social science disciplines investigating factors that impede and/or trigger its occurrence. We draw upon two specific strands of collective action literature to derive a set of key features which are generally associated with a higher level of collective action. We confront these key features with empirical evidence from our Ugandan village setting and conclude that most of the key features are present for collective action to thrive. Yet, we simultaneously observe little instances of community monitoring in reality which puts each of the distinct theoretical approaches into perspective. Our article demonstrates that it is only the intersection of both theoretical perspectives that furthers our understanding of the paucity of community-based monitoring.

Our study is not only timely and topically relevant for an (monitoring and) evaluation audience, also the use of social network analysis which we consider an innovative evaluation method is interesting for an EES audience.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 061 Participatory and Stakeholder Approaches

## O 186 - Formative Participatory Evaluation Critical for Quality Assurance in African Universities

*O. John Michael Maxel<sup>1</sup>*

<sup>1</sup> Uganda Evaluation Association, member, Kampala, Uganda

**Rationale:** Quality assurance in African Universities has taken a dynamic turn more using participatory approaches. Traditionally universities in Africa used to undertake an evaluation of the programmes accreditation bodies looking at the systems. The traditional method of quality assurance in universities mainly with a focus of improvement in the university system, process, procedures and mainstreaming actions intended for the achievement, maintenance, monitoring and achievement of quality in the University. African Universities and quality assurance bodies have undertaken similar paths putting in place standards, grades, classes, status, conditions, character building and value building mechanisms.

Despite the efforts, traditional quality assurance systems remained inadequate in Africa while transformative participatory evaluation has become critical in the evaluation of university programmes. Transformative participatory evaluation implores principles of participatory evaluation which envisages democratizing social change but with different ideological and historical roots (Cousins & Whitmore, 1995). This approach primarily emerged in developing countries (Fals-Borda, 1980; Fernandes&Tandon, 1981; Tandon, 1981; Kassam& Mustafa, 1982) in the development assistance programmes has become critical in the delivery services. Traditional quality assurance mechanisms mainly derived from positivism paradigm detaching participants therefore poorly addressing social and economic problems. Cousins and Whitmore (1998) argue that T-PE now spread to the university sector.

**Purpose:** The purpose of this study is to provide the impact of insight of transformative evaluation in University.

**Justification:** The paper provides insight into the dynamics which have been used by some Universities in Africa to undertake the participatory evaluation. Work environment, service evaluation, faculty evaluation, systems evaluation, process evaluation by the users and stakeholders involved in the delivery of services in the University systematically transformed the delivery of the services in Mount Kenya University. This systematically improved the by using transformative participatory monitoring and evaluation is empowerment of people (Cousins & Whitmore, 1998), constructing and respecting participants own knowledge and understanding, power and control (Fals-Borda & Anisur-Rahman, 1991; Tandon, 1981). TPE enabled educators, students, administrators and evaluators to interact and improve approaches for improvement in the systems and process in the university collectively. Rather than police and criminal perspective TPE rather education process that involves and systematically transforming the entire university community. TPE provides a mechanism for the involvement of students, faculty, and entire University community as their interests, preoccupations, aspirations, and priorities are taken care off (Brunner & Guzman, 1989).

**Methodology:** The paper used literature review a means for primary data collection together qualitative and quantitative methodologies for data collection from the different practitioners in monitoring evaluation on negative vice slowly destroying fabrics of monitoring and evaluation in Africa.

**Conclusion:** Transformative evaluation is critical in an enhancement of the practice of monitoring and evaluation practice in different organizations. In order to improve the quality of monitoring and evaluation practice, the strategy is crucial to the advancement of practice in dynamic global environments.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 062 Improving Systems and Capacities Through Evaluation

## O 188 - Towards a General Theory of Monitoring and Evaluation: A Method to Balance Donor Budgets, Agency Capacity and Project Impact

*E.E. Goetsch*<sup>1</sup>

<sup>1</sup> *University of the Witwatersrand, School of Law, Johannesburg, South Africa*

**Rationale:** In times of austerity, donors cut budgets. When they do, the capacity of agencies and the impact of their projects shrink as well. Capacity is key to impact. It starts the project cycle (input->throughput->output->outcome) and the results-chain (capacity->performance->delivery->impact). When proposals are submitted and contracts are negotiated, donors need to see and agencies need to set the boundary between fat and muscle. When M&E practitioners write reports, we want a robust method to measure the capacity gaps and explain or signal project failure. This presentation suggests a method to link budget, capacity and impact using the organogram. Without such a link, donors can demand more than they pay for and agencies can accept a burden they cannot carry.

**Objectives sought:** To present a method to link budgets, capacity and impact that can credibly serve the profession as a standard tool for right-sizing project budgets and impact projections. In this way, to increase the aid-effectiveness of donors and cost-effectiveness of agencies.

**Brief narrative:** The method combines a number of disciplines. It uses management information from Finance, Human Resources, Operations and Strategy to compare the organisation's supply of labour with the project's demand for labour. The supply is the sum of directors, managers and workers in position and the demand is the product of governance units and key performance areas. The method has dramatic results. The average capacity of World-Bank assisted projects in oil-rich Nigeria was found to be 20% of requirements. This figure is in line with the 30 implementation agencies evaluated in March 2018.

**Justification:** The method offers a standard for triangulating budgets, capacity and impact. It has many benefits. It rescues agencies from "extreme haircuts". It helps evaluations help people to improve their lives and make our societies more resilient. It reduces unpredictability and complexity. It helps when designing and managing evaluations. It informs evaluation systems. It helps to rethink evaluation methods, design and criteria. It combines methods in evaluation. It addresses delivery risk in uncertain futures. It facilitates collecting and analysing data and reporting issues particularly in challenging contexts. It provides a dashboard that integrates ICT, M&E and managers. It enables evaluation to become foresight. It rewards flexibility and handles complexity. It develops the field of Evaluation to promote resilience and action in this critical time. It solves a challenge and offers opportunities for the evaluation field. It addresses some of the dilemmas and trends in professionalism, standards and ethical norms. It advances the theory and applies ethical values to evaluation. It promotes Evaluation Associations as custodians of a professional standard. It protects the independence of evaluators and our relevance and responsiveness. It supports the partnerships and stakeholders who make up the international development aid industry and greatly strengthens the communicating, using and embedding of evaluation. It is truly intersectional, since it connects M&E to every role-player in the organisation.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 062 Improving Systems and Capacities Through Evaluation****O 189 - Examining the Extent and Effects of Disability-Sensitive Community Monitoring on Educational Service Delivery in Katakwi District, Uganda**L. Popelier<sup>1</sup><sup>1</sup> University of Antwerp, Institute of Development Policy, Antwerpen, Belgium

Despite the formulations of conventions, laws, policies and programs at various levels related to the provision of inclusive services, the implementation is often hindered by persistent supply side governance constraints. However, failures in public service delivery are also increasingly seen as failure in accountability relations (e.g. World Bank, 2004). Strengthening public accountability has, therefore, gained importance as a key strategy to enhance public service delivery (Joshi, 2013). The adoption of bottom-up accountability activities, such as community-based monitoring, has inspired many researchers to examine whether and under what conditions such activities can achieve intended results, particularly improved quality of local services (Brinkerhoff & Wetterberg, 2016; Dewachter & Holvoet, 2017; Holland, Ruedin, Scott-Villiers, & Sheppard, 2012; Joshi, 2013). What is still largely lacking in this debate, however, is a more in-depth discussion on whether community participation in accountability mechanisms automatically results in socially inclusive and equitable service delivery. In other words, the heterogeneity of and power distributions among right holders are often ignored. This may have major consequences, given that the likelihood that bottom-up accountability mechanisms result in equitable service delivery depends on the extent to which the involved citizens represent or negotiate on behalf of all citizens, including worst-off groups such as persons with disability (Fox, 2014). Therefore, this paper first examines the extent to which perspectives and needs of persons with disability are considered in the community-based monitoring activities conducted in three parishes in Katakwi district, Uganda. Subsequently, the researcher draws upon the theory of evaluation influence (see Mark & Henry, 2004) to reflect upon the effects of a capacity building intervention which aimed to foster the incorporation of a disability-lens in the community-based monitoring activities. More precisely, attention is paid to the effects of this capacity building on the future monitoring activities (e.g. monitoring tools), the responsiveness of the monitoring outputs to the needs of persons with disability and the equitability of the monitored primary schools. Intentional variations in the type of stakeholders who were engaged in these capacity building interventions in each of the three targeted localities also allow for a reflection on the effect of stakeholder selection and representation on the monitoring activities, outputs and outcomes.

Building resilient societies will require bottom-up accountability mechanisms that involve and represent all citizens, including persons with disability. The way in which persons with disability experience and are affected by policies, projects and programs often significantly differs from the experience of able-bodied persons. Nonetheless, monitoring and evaluation practices rarely take into consideration their perspectives. This study aims to raise awareness on the potential consequences of the continuous disregard of the perspectives of persons with disability in M&E, including community-driven forms of M&E. Yet, the paper equally aims to highlight potential benefits in terms of evaluation outputs and outcomes of adopting a disability-lens in community-based monitoring activities.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 063 Using Evaluation To Build System Capacity

## O 190 - Integrating Social and Behavior Change Communication Indicators into a National Health Information Management System: The Ghana Experience

*Y. Ampeh<sup>1</sup>, G. Annan<sup>1</sup>, A. Ofosu<sup>2</sup>, E. Sefa<sup>3</sup>, D. Atweam<sup>4</sup>, C. Niikoi<sup>5</sup>*

<sup>1</sup> Ghana Health Service, Health Promotion Department, Accra, Ghana

<sup>2</sup> Ghana Health Service, Policy- Planning- Monitoring and Evaluation Department, Accra, Ghana

<sup>3</sup> FHI 360/Ghana, Communicate 4 Health, Accra, Ghana

<sup>4</sup> Ghana Health Service, Policy- Planning- Monitoring and Evaluation, Accra, Ghana

<sup>5</sup> UNICEF-Communication 4 Development, Communication 4 Development, Accra, Ghana

**Background:** Very few Ministries of Health and departments have comprehensively developed and incorporated M&E systems and social and behavior change communication (SBCC) indicators into their health management information systems. A lack of a robust M&E system and SBCC indicators for measuring Ghana Health Service Health Promotion Department (GHS/HPD) interventions was identified by the department as a major challenge for several years. This negatively impacted SBCC data capture, analysis, reporting and the profile of GHS/HPD nationwide. In 2015, Health Promotion Department of the Ghana Health Service as part of its core mandate of strengthening the capacity of HPD staff to coordinate and deliver evidence based SBCC interventions, led the path to develop a robust M&E system and SBCC indicators for HP. The department using its excellent networks and leadership skills brought together a consortium of partners who provided technical assistance and/or funding to include SBCC indicators into DHIMS2.

**Description of Intervention:** A consortium of partners comprising HPD/GHS, C4H, UNICEF, Policy Planning, Monitoring, and Evaluation Division/GHS, Family Health Division / GHS and USAID implementing partners (Evaluate for Health and Systems for Health) formed a 15-member technical working group which developed a road map for a robust M&E system for HP. Capping a period of seven months the group held several meetings and brainstorming sessions which resulted in the development of a results framework and indicator table; collation, review and revision of existing data collection and reporting forms; development of new tools and reporting forms as needed and accompanying data dictionaries; stakeholder engagements which reviewed and provided needed inputs; and pretesting and finalized tool/forms. These led eventually to a “Boot Camp” in late 2015 where SBCC indicators were eventually incorporated into DHIMS2 Platform with co-funding by C4H and UNICEF.

**Results:** In total 33 SBCC indicators were developed, 22 of which are captured routinely by health staff at various levels and entered into the DHIMS2 platform. Eight primary data collection tools (all paper-based) and three reporting forms were also developed. Both C4H and UNICEF supported the training of all regional and district officers across the entire country on the use of the tools. Over 12,000 data collection registers were printed and distributed throughout the country and many copies are required to sustain continued data collection. Printing paper-based registers comes with a huge burden that governments and partners might not be able to sustain.

**Implications for the field:** Since August 2017, SBCC indicators have been captured into DHIMS2 with close to 100% reporting rates across the entire nation. Collaboration among partners has been key to the development and implementation of a robust M&E system for HP as partners brought on board diverse technical expertise and resources. GHS/HPD capacity was built through a “learning by doing model” with its involvement right from design to implementation of the M&E system. A sustainability plan ensuring continued reporting of SBCC indicators into DHIMS2 and use of data for decision making has been drafted. Core to this plan is transitioning from paper-based to electronic data collection.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 063 Using Evaluation To Build System Capacity**

**O 191 - Learning for Adapting: Lactational Amenorrhea Method (LAM) for Family Planning in the Philippines**

*E. Bautista<sup>1</sup>, R. Flueckiger<sup>2</sup>, E. Dasmarias<sup>1</sup>, R. Colaco<sup>3</sup>, R. Benabaye<sup>1</sup>, F.A.R. Gillego<sup>4</sup>, A.M. Lozada<sup>6</sup>, C. Bisson<sup>3</sup>, M. Chen<sup>3</sup>*

<sup>1</sup> RTI International, LuzonHealth Project, Pasig City, Philippines

<sup>2</sup> RTI International, International Development Group- Global Health Division, Research Triangle Park, USA

<sup>3</sup> RTI International, International Development Group- Global Health Division, Washington- DC, USA

<sup>4</sup> Legaspi City, Health Officer, Legaspi City, Philippines

<sup>5</sup> United States Agency for International Development, Philippines, Manila, Philippines

<sup>6</sup> Legaspi City, Department of Health Regional Office V., Legaspi City, Philippines

The World Leaders Declaration on Population states that (1) the population problem is a principle element in long range national planning, (2) the opportunities to decide on the spacing and number of children is a basic human right, (3) peace depends on how population growth is met, and (4) the objective of family planning is the enrichment of human life and not its restriction.

Family planning is a key issue in the Philippines, where the 2013 National Demographic and Health Survey (NDHS) estimates 54% of married women 15–49-years-old do not want another child and 19% want another child but prefer to wait two or more years. Yet, only 38% of married women report using a modern method of contraception. A reported 11% of pregnancies are unwanted and 17% are mistimed. Importantly, 26% of births occur less than 24-months after previous births. Highlighting the need for family planning education and intervention.

Lactational Amenorrhea Method (LAM) for family planning is at least 98% effective at preventing pregnancy when all criteria are met; (1) the mother is exclusively or nearly exclusively breastfeeding, (2) the mother is amenorrhoeic, and (3) the baby is six months old or younger. LAM is a common and accepted form of family planning in the Philippines.

The USAID funded, LuzonHealth project conducted a prospective evaluation in Legaspi City to inform the enhancement of guidelines aimed at increasing LAM compliance and encouraging a second form of contraceptive once LAM protection expires. LAM compliance, reasons for non-compliance, family planning referral and uptake of secondary modern family planning methods were tracked over a nine-month period among 521 postpartum women.

Evidence-based learning is critical for effective adaptation of project implementation. The use of appropriately designed operational research is critical for assessing the successes and limitations of project implementation, that cannot be measuring using planned monitoring and evaluation (M&E). Operational research provides defensible evidence for decision-making, strengthening the ability for projects to adapt. The provision of clear and statistically sound results serves as a catalyst for policy change at many levels. Outcomes from this evaluation have been used to inform postnatal follow-up guidelines and referral practices and encourage community health workers to track LAM compliance during follow-up visits.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 063 Using Evaluation To Build System Capacity**

**O 192 - Improving Public Health Management Advances Resilient Societies: Evaluation Findings of a Workforce Capacity Building Program in Kenya and Bangladesh**

*C. Matson<sup>1</sup>, E. Willacy<sup>1</sup>, A. Yoos<sup>1</sup>, P. Wheeler<sup>1</sup>*

<sup>1</sup> *Centers for Disease Control and Prevention, Division of Global Health Protection, Atlanta, USA*

How does a public health management training program build resilient societies? Public health managers are essential to health systems – detecting, assessing, reporting, and responding to public health threats. Yet for low and middle-income countries, few managers have the skills or training to manage public health crises. For example, the 2014–2016 Ebola outbreak exposed mismanagement in logistics and supply, commodity management, coordination between local and international public health organizations, planning for staffing and financial concerns, and engaging local communities.

With the Ebola experience in mind, the Centers for Disease Control and Prevention (CDC) designed Improving Public Health Management for Action (IMPACT). IMPACT is a workforce capacity-building program that develops highly-skilled public health managers. These managers will be able to translate the science of public health into action, improving the health of their local communities, and thereby building more resilient societies. Launched in 2016 in Bangladesh and Kenya, the program recruits local Ministry of Health staff into a service-based training program ranging from from six months to two years, depending on participants' seniority and experience.

Since inception, evaluation has been an integral part of IMPACT's planning and execution. First, IMPACT fellows receive education in evaluation theory and methods. Second, the IMPACT evaluation team regularly evaluates the process, outcome, and impact of the program. Evaluation results are used for two purposes: to improve the training delivery and outcomes for future cohorts; and to encourage more investment in public health management training worldwide.

To meet these purposes, we have collected data on the program and its fellows, using the Success Case Method, surveys, semi-structured interviews, and site visits. In this paper, we describe the achievements of our cohorts quantitatively, based on an assessment and analysis of their progression in eight management competencies. Then, we explore the program's achievements qualitatively through four fellow case studies. Each demonstrates the impact of improved management in a distinct area: detection and resolution of flaws in surveillance systems and commodity management; an assessment of a community's health needs and assets; delivery of effective risk communications; and management changes to improve response to a refugee crisis.

These evaluation findings demonstrate how IMPACT fellows have contributed to protecting and preventing emerging public health threats in Kenya and Bangladesh. Moreover, the findings emphasize that to build resilient societies, countries must have the ability to galvanize organized action to combat public health threats.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 005 Institutionalising Evaluation**

**O 193 - The Evolution of Spatial Operational Programmes' Evaluation in Greece: A Market Approach**

*N. Koutsomarkos<sup>1</sup>, G. Georgiadis<sup>1</sup>*

<sup>1</sup> Aristotle University of Thessaloniki, Department of Spatial Planning and Development, Thessaloniki, Greece

The objective of this paper is to map the market of evaluation studies of the NSRF 2007–2013 and the 4<sup>th</sup> Programming Period until now and particularly in the regional and territorial cooperation scale, in order to provide insights and recommendation for the improvement of the evaluation ecosystem in Greece.

The culture of programme evaluation in Greece goes hand to hand with the requirements of the European Structural Funds regulations since the 2000–2006 Programming Period. The market and the profession has matured along the progress of the evaluation specifications, but evidence suggests that the employment of supplementary and/or innovative approaches and methods are limited.

What is the role of the market size in the evolution of the profession? What are the drives for better evaluation approaches?

A systematic analysis of the institutional evaluation framework in Greece, regarding ROPs and ETC, builds the base for the above questions. Then the market for the said evaluation is mapped out in terms of size, volume and players. A multitude of evaluation studies deliverables are examined and classified in categories highlighting the tools and methods employed. The empirical insights of authors professional experience supported by the data analysis suggest that the market of evaluation studies in Greece will be improved by facilitating better competition on qualitative terms, lower entrance hurdles for new approaches and players, a broader allocation of financial resources from the interested public organizations and an improved management of evaluation contracts and results.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 065 Designing Development For Change

## O 194 - Making Evaluations Matter: Lessons Learnt from a Reorientation Process at the Media Development Organization DW Akademie

P. Berner<sup>1</sup>, J. Lublinski<sup>1</sup>, D. Reineck<sup>1</sup>

<sup>1</sup> DW Akademie, Strategy and Consulting Services, Berlin, Germany

One of the challenges in evaluation is getting them to effect positive change. All too often, evaluation results end up in drawers, with people returning to business as usual once the dust has settled. This is unfortunate, since the ultimate purpose of evaluations is social betterment (see Mark/Henry/Julnes 2000). Without adequate reflection, opportunities to make improvements are neglected. The aim of the presentation is to share solutions from our own organization, DW Akademie, with the evaluation community and to spark a conversation on what strategies might be implemented to maximize the effectiveness and impact of formative evaluations in an organizational context.

The presentation will comprise of three parts, drawn from our experience in moving from an organization engaged in short-term journalism training projects to a full-fledged media development organization running long-term projects:

1. Inclusion of evaluation stakeholders;
2. Improvements by means of dedicated systematic communication;
3. Improvements by means of organizational processes.

In the first part, we identify the different stakeholders in the evaluation process, focusing our attention especially on the perspectives of project management and evaluation department (while also taking into consideration donors, partners and target groups). This is important, since these two parties often have conflicting expectations from the evaluation (see Bryson, Patton, Bowman 2011). And they tend to blame one another when evaluation results do not lead to change: Managers criticize evaluation reports for being too general in focus and too far away from their everyday work. Evaluators tend to blame managers' unwillingness to acknowledge their mistakes. Including all stakeholders and taking into account their expectations is key to ensuring evaluations have an impact.

The second part will concentrate on dedicated communication as a means of improving uptake of evaluation recommendations. This includes methods not recommended by ourselves (external pressure, decommissioning of management), as well as methods that we consider to be good practice (advance briefings, participatory planning, change workshops). We share how and through which formats project management is informed and involved in the run-up to and the analysis of evaluation results at DW Akademie.

In the third part, we focus on organizational processes implemented to maximize the usefulness of evaluations. Here too, we discuss methods rejected by ourselves (trial type hearings, hierarchical decisions), but focus on good practice (recommendations-guided strategy development, linking evaluation with consultancy, process and time planning at different levels). DW Akademie has developed a learning cycle that interlinks the evaluation and strategy development phases of 25 BMZ-funded country projects in three-year-project cycles. One main objective here is to create space and opportunities for learning from the past (diachronic) as well as from projects taking place elsewhere (synchronic). The evaluations also focus on digital aspects of the projects, taking into account the tectonic shifts currently taking place in the media landscape.

Our hope is that our presentation will be the starting point of a conversation and exchange between session participants, enabling all to learn from other organizations on how they facilitate and integrate learning and positive change, drawn from the insights they have gained via monitoring and evaluation.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 065 Designing Development For Change

## O 195 - Challenges Evaluating New Approaches to Programme Design: Evaluating Human-Centred Design for Behaviour Change in Adolescent Sexual and Reproductive Health

S. Wallach<sup>1</sup>

<sup>1</sup> Itad, Organisational Effectiveness, Hove, United Kingdom

As funders and implementers recognize that 'business as usual' is not working to address complex problems, they are testing new approaches to program design, and are eager for robust evaluation to inform learning. Human-centred design (HCD), an iterative process of formative research and testing, prototyping and piloting interventions with end-users, is gaining interest as an approach. This new approach has posed challenges in evaluation design and implementation including:

- Uncertainty around key information up-front (targets, geographies, the intervention) make the evaluation design resource intensive as it needs to be revisited as the intervention is developed.
- In early applications of HCD, a lack of formal documentation made it difficult to track the HCD process.

Alongside this, there has been a heightened need for robust evaluation to justify investments in a new approach. These challenges necessitate an adaptive M&E approach with close engagement with evaluation partners. This session will explore some of the ways to address challenges evaluating an HCD approach to program design. It draws on Itad's experience evaluating two HCD interventions which aim to increase contraceptive use among adolescent girls:

1. Adolescents 360 (A360), a 4.5 year initiative funded by Bill and Melinda Gates Foundation and Children's Investment Fund Foundation (2016–2020) implemented in Ethiopia, Nigeria and Tanzania.  
A360 includes a process evaluation (PE), outcome evaluation (OE), and cost-effectiveness study (CES). It aims to understand how and if A360 leads to better SRH outcomes, generate evidence about A360's cost-effectiveness; understand implementation in different contexts and support adaptive management.
2. The Hewlett Foundation's strategy to apply HCD to ASRH through funding IDEO.org and Marie Stopes in Kenya and Zambia. This evaluation (concluded), drew on qualitative methodologies (e.g. focus group discussion, key informant interviews) and various analytical methods to test a theory of change for HCD to understand the potential and limitations of HCD to designing interventions for adolescent girls.

HCD has the potential to enhance ASRH programme impact, yet measuring its application and impact pose some challenges due to the extended, iterative design process. As a new approach to program design, there has been a call for robust evaluation of HCD from funders and implementers to provide decision-makers with the information they need to make informed decisions about the value of the approach, to support learning about what does and doesn't work and to enable implementers to adjust their approach and course correct. In this space, evaluation needs to be rigorous and flexible, and conducted in a way that builds credibility of the evaluation and ownership of the findings by funders and implementers.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 066 Complexity and Causal Modeling

## O 196 - Emergent Design of the Ecosystem Services for Poverty Alleviation (ESPA) End of Programme Review

*J. Colvin<sup>1</sup>, M. Wells<sup>2</sup>, A. Moore<sup>3</sup>, U. Pascual<sup>4</sup>*

<sup>1</sup> Emerald Network Ltd, MLE, Stroud, United Kingdom

<sup>2</sup> Michael Wells Associates, Evaluation, Oslo, Norway

<sup>3</sup> Palladium International Ltd, Evaluation, Bristol, United Kingdom

<sup>4</sup> Basque Foundation for Science, Basque Centre for Climate Change BC3, Leioa, Spain

The turbulent times we are now living through are characterised by increasing complexity and uncertainty, calling for new approaches to resilient development and evaluation methodologies that can match these. Working with complex situations requires us to attend to complexity-sensitizing concepts such as non-linearity, dynamic interactions, coevolution, adaptation and emergence. These concepts also require consideration within evaluation practice, leading to innovations – such as developmental evaluation – in both formative and summative contexts.

In this paper we consider three strands of innovation in complexity-based evaluation that is designed to promote resilience. These are: adaptive governance of social-ecological systems as foundations for resilience; principles-based evaluation; and co-producing emergent design. To illuminate these strands we reflect on the emergent design of the end-of-programme review of the 10-year Ecosystem Services for Poverty Alleviation (ESPA) research programme, which a team of four of us undertook in late 2017.

While the framing of the review by the client (DFID) focused on specification of review questions, methodological requirements were permissive, enabling us to explore an approach which maximised contextual fit with the subject matter of the review, namely the governance of the ESPA programme, its research findings and development impact.

In terms of the adaptive governance of a social-ecological system, the review cast light on both the challenges and successes of designing and governing an effective, global research-for-development-impact programme in this complex area which necessarily (given the ecological underpinnings of human wellbeing) lies at the heart of efforts towards more resilient development. Major findings concerned the self-organising dynamics of the programme, and the relative contributions and alignment of top down governance and bottom up innovations in inter and transdisciplinary research. Revealing these dynamics required a mirroring in the emergent design practice of the evaluation, progressively illuminating a 'muddled middle' by drawing on deductive, inductive and abductive reasoning.

Emergent design practice is guided by the principle that in complex systems, design can only ever be fully understood in retrospect – hence emergent. To put this another way, starting with a fully formed methodology without fully understanding and agreeing the system of interest would be maladaptive. Even in a relatively rapid summative review it makes better sense to hold open the space for iterative, collaborative learning about the system of interest for as long as possible, and then to synthesise the emerging picture in a coherent argument. In the ESPA review, the four of us had not previously worked together, requiring an attitude of trust and critical appreciation in support of emergent design.

Our design praxis also highlights the value of a principles-based approach, with a nested set of principles guiding the review and facilitating trust-based coproduction. Beneath the overarching principle of a contextually-responsive approach lay a cluster of systems-based principles – diverse, multi-voiced, iterative and emergent, reflexive and illuminating the muddled middle.

The paper finishes with a consideration of the wider significance of these three strands of innovation in the context of rethinking evaluation methodology to deal with flexibility and complexity, and to promote resilient action in critical times.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

S 066 Complexity and Causal Modeling

## O 197 - Economic Multiplier of an Asset Transfer Model: The Case of Heifer Malawi

E. Fitzpatrick<sup>1</sup>, S. Akgüngör<sup>2</sup>

<sup>1</sup> University of Arkansas, Clinton School of Public Service, Little Rock, USA

<sup>2</sup> Dokuz Eylul University, Economics, İzmir, Turkey

**Rationale:** Development programs that aim to enhance the welfare of a community have a target population which is usually a subset of a local economy. The population that surrounds and interact with the target population may also enjoy benefits of the intervention through spillover effects. Spillovers are secondary effects of interventions where non-participants gain from the economic change stimulated by the intervention. Examination of spillovers allows us to quantify the degree of interconnectedness that evolves from project activities and from this interconnectedness, the likelihood of more resilient communities with sustainable impacts.

**Objectives:** This study examines the extent to which indirect effects or spillovers result from an asset transfer and capacity building intervention implemented by Heifer International in Malawi (HI Malawi). The data consists of a baseline survey in 2014 and end of project survey in 2017. We selected two groups of farmers: treatment group and a comparison group. The comparison group was statistically matched and geographically separate from the treatment group. In both years, we conducted a social network survey to reveal the network ties of treatment and comparison groups and created four samples: treatment group, control group, treatment network and comparison network. We compared the change in net farm income of the participants, their non-participant network and the comparison group capturing the spillover effects as an elasticity estimate. This measure is the ratio between "percentage change in the value of the variable in the treatment network" and "percentage change in the value of the variable in the treatment group" with respect to the comparison group. We found that participant economic activity had a significant impact on the net income of non-participants due to the creation of backward and forward linkages and knowledge exchange.

**Narrative and justification:** The motivation of this study is to develop and discuss the use of a field-based methodology for assessing the total economic effect of an intervention by including the direct and indirect effects of the asset and capabilities transfer. The paper proposes a method to capture the spillover benefits of a project so that the evaluation would be more comprehensive. Previous evaluation results reveal that rural livelihoods (operationalized by net farm income, social capital and women empowerment index) has increased with the HI Malawi project. In this study, we further explore whether improvements in rural livelihoods were seen in the population that interacts with the target population, tracing out the indirect effects of backward and forward linkages. The technique that the paper develops and demonstrates allows an evaluation to systematically capture the catalyzing effects of program intervention beyond the direct effects, providing a more complete estimate of total program impact.

Thursday, 4 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 067 Emerging Evaluators: The Next Generation**

**O 198 - Quick Guide to Incorporate Young and Emerging Evaluators to Organizations: What Works?**

*C. Olavarria<sup>1</sup>, G. Sánchez Romero<sup>2</sup>, A. Guidoccio<sup>3</sup>*

<sup>1</sup> EvalYouth LAC, Research, Santiago, Chile

<sup>2</sup> EvalYouth LAC, Research, Mexico City, Mexico

<sup>3</sup> EvalYouth, Global, Buenos Aires, Argentina

Since year 2015 a strong movement of young and emerging evaluators is arising in Latin America with a demand for a space to take part in the strengthening processes that are taking place in the evaluation field in the region.

Young and emerging evaluators are building regional and national chapters of EvalYouth and building strategies to get involved in the process of achieving the EvalAgenda 2020 goals with great commitment. However, a close involvement of these young evaluators teams with regional and national evaluators organizations is required to put this energy into value and ensure the development of all its potential.

The quick guide presented in this paper was developed as a compilation of lessons learned from the experience of 7 national evaluation associations and 1 regional association in Latin America. It proposes a series of steps and tips to successfully integrate young evaluators to organizations in the LAC region. Likewise, this is particularly relevant to African, Asian and Arab countries as many EvalYouth local and regional chapters are being created.

Youth and young evaluators are key stakeholders when analyzing the sustainability of the evaluation field worldwide and especially for the strengthening of Voluntary Organizations of Professional Evaluators (VOPE).

Within the Latin-American experience, the findings show the great potential that incorporating young evaluators has for the VOPE teams in terms of innovation, bringing new ideas, use of technologies, embracing global approaches and new methods, developing commitment with the evaluation field and promoting cross-generational thinking.

This quick guide provides a space for debating lessons learned from the Latin American case and envision how these processes can be useful in other regions of the world.

Thursday, 4 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 067 Emerging Evaluators: The Next Generation**

**O 199 - The EvalYouth National Chapters in LAC: a Comparative Study on the Strategies of Conformation and Sustainability of Voluntary Action**

*V. Terra Polanco<sup>1</sup>, J. Arévalo<sup>2</sup>, M. Ordóñez<sup>3</sup>*

<sup>1</sup> *Universidad de Chile, Departamento de Antropología, Santiago, Chile*

<sup>2</sup> *National Institute of Educational Evaluation, Direction, Quito, Ecuador*

<sup>3</sup> *External consultant, Public policies, Ciudad de México, Mexico*

The main purpose of this paper is to open up the opportunity to share successful experiences and lessons learned from seven Latin American countries that have created national chapters of EvalYouth LAC, in their effort to promote and build capacities of young and emerging evaluators (YEEs). On the one hand, the paper will promote a space of reflection and discussion about lessons and guidelines that help build common agendas to strengthen the networks of young and emerging evaluators in the region. On the other hand, the discussion will contribute to the sustainability of the EvalPartners' network, EvalYouth, by highlighting the maintained contributions that young and emerging evaluators can make in the field of evaluation.

Accordingly, the paper will present the findings of a comparative study carried out by EvalYouth LAC in 2017. The study addressed the different strategies developed to create the national chapters of EvalYouth in Argentina, Bolivia, Chile, Costa Rica, Ecuador, Mexico and Paraguay during 2017. It also explored both the strategic plans and actions these countries developed to put into practice the global initiative.

In the first section, the paper will show the strategies of conformation in each country, delving in the singularities, differences and results obtained. Then, in the second part, it will present an overview of the strategic plans and actions carried out by the different countries mentioned above. This overview will focus on common points and complementarities among these countries, as well as the challenges and lessons learned in the process of conformation of local chapters.

Finally, the paper will conclude by presenting a summary of the different strategies, methodologies and partnerships to promote the building, development and strengthening of evaluative teams; especially the networks of young and emerging evaluators. The latest by identifying successful practices, lessons learned, as well as critical points.

Thursday, 4 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 067 Emerging Evaluators: The Next Generation**

**O 201 - Young and Emerging Evaluators in Latin America:  
Their Position and Role in the Evaluation Field**

*V. Terra Polanco<sup>1</sup>, M. Ordóñez<sup>2</sup>*

<sup>1</sup> *Universidad de Chile, Departamento de Antropología, Santiago, Chile*

<sup>2</sup> *External consultant, Public policies, Ciudad de México, Mexico*

The paper shows the results of six surveys on the situation and role of the young and emerging evaluators (YEEs) in Latin America and the Caribbean region. The study was carried out in 2017 by members of local chapters of EvalYouth in the following six countries of the region: Argentina, Bolivia, Costa Rica, Chile, Ecuador and Mexico.

First, the paper will present the methodology used to develop the study. The six countries made use of an online questionnaire that inquired several issues related to the level of involvement, relevance and incorporation of young and emerging evaluators in the evaluative community in each country. The questionnaire was completed by young evaluators that work or study in public and private institutions, such as universities, civil society organizations, consultancies and public services, among others.

Then, the presentation will show the situation of young and emerging evaluators, addressing variables such as age, gender, study, type of work, evaluation experience, expectations and perspectives. It will also draw on the role of these young evaluators, exploring their involvement in evaluation teams, their involvement in local volunteer organizations for professional evaluation (VOPEs), the value of the role and contribution they can make to the evaluation field, and the training offer in evaluation, among others. The focus will be placed on the similarities and differences among the six countries, in order to generate recommendations and improvement proposals to strengthen the capacities and protagonism of young and emerging evaluators within the region.

The information obtained in the diagnosis will be useful to define the main guidelines of a Latin American and Caribbean agenda. In particular regarding to the visibility and strengthening of the young and emerging evaluators in the region.

Thursday, 4 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 068 Using Social Network Analysis for Evaluation Purposes****O 202 - A Scoping Review on the Current and Potential Use of Social Network Analysis for Evaluation Purposes***L. Popelier<sup>1</sup>*<sup>1</sup> *University of Antwerp, Institute of Development Policy, Antwerpen, Belgium*

Increased recognition of the omnipresence and importance of relationships and networks has fueled the development of the social network analysis (SNA) approach, which considers structural relationships to be primary causes of societal outcomes. While the potential of SNA has been demonstrated and discussed extensively in social science research, interest in applying it in the field of evaluation has only emerged recently (Laven et al., 2010; Durland and Fredericks, 2005a). This evolution can be explained in part by the fact that the quality and effectiveness of a given intervention are increasingly being linked to the structure and relational quality of interactions among the stakeholders involved in that intervention (Abma, 2006; Davies, 2003; Giuffre, 2013; Gomersall et al., 2017; Hunter et al., 2015). This scoping review of journal articles therefore examines the current and potential use of SNA for evaluation purposes and reveals that evaluators use SNA because of its ability to identify key stakeholders, assess network structures and relationships quantitatively, reveal informal relations and visualize even complex networks. However, challenges arise when interpreting findings, determining causation between network structures and outcomes and disseminating evaluation results in an ethically responsible manner. The review concludes that the evaluation field – especially in the development sector – would benefit greatly from increased use of SNA, but that this would first require improved use of alternative sources of network data, qualitative methods and inferential statistics that will enable evaluators to move beyond descriptive network analysis.

The content of the submitted abstract has been presented in a poster format during the AfrEA conference in March 2017. The presented poster was selected as one of the three best posters and therefore received an award from the editor of the African Evaluation Journal who praised the work because it provides a useful overview of the opportunities and the potential pitfalls of an innovative methodological approach within the evaluation field. The paper serves as a good introduction for people who are rather unfamiliar with SNA approaches, because it presents a framework of evaluation questions that could be addressed and discusses the possibilities in terms of data collection and analysis. Yet, the critical discussion of the strengths, weaknesses, opportunities and risks of the approach presented in the paper is also likely to generate fruitful debate among members of the evaluation community who are more familiar with the SNA approach. In addition, the further exploration of approaches such as SNA which are able to describe and explain relations and structures will also assist in the strive towards more resilient societies given the increased recognition that social arrangements can affect collective and societal outcomes.

Thursday, 4 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 068 Using Social Network Analysis for Evaluation Purposes**

## **O 203 - Using Social Network Analysis to Evaluate the Network Formation Among Fellows in a South African Leadership Development Programme**

*J. Govender<sup>1</sup>, M. Geeta<sup>1</sup>, W. Church<sup>2</sup>, B. Klugman<sup>3</sup>*

<sup>1</sup> *Data Innovator, Johannesburg, South Africa*

<sup>2</sup> *Independent, Seattle, USA*

<sup>3</sup> *Barbara Klugman Concepts, Johannesburg, South Africa*

**Rationale:** This presentation provides the findings of Stage 1 of the social network analysis (SNA) that forms one component of a broader developmental evaluation commissioned by Atlantic Philanthropies for the Tekano Fellowship Programme. Tekano's mission is "to foster dynamic, visionary, values-based leaders working both individually and in catalytic communities of learning and action who articulate, convey, and act to promote and achieve health equity." The evaluation findings will be used to inform Tekano's continuous improvement of the fellowship programme as well as to inform its funders.

The fellowship's theory of change includes the assumption that the programme experience will contribute towards building trust and fostering a catalytic community networking and advocating for health equity. Tracking and visualizing network formation over time allows the Tekano programme developers to identify networking trends and adapt the programme to further facilitate networking among fellows. Hence its value for the utilization-focused approach of this evaluation.

Social network analysis (SNA) is a method and set of metrics to measure, visualize and draw inferences on the characteristics of a group or groups of individuals. It has been used for various kinds of networks but is particularly useful in evaluating leadership networks. The SNA data collected for this study was analyzed using an open-source software programme called "Gephi."

**Objectives sought (primary question):** Are the fellows – as individuals and as the groups that some applied in – developing as a network that supports each other's activities in advocacy for social determinants of health equity?

**Narrative and justification:** Quantitative SNA data as well as some qualitative data was collected using an online survey created in Google Forms. The form was administered to the 34 fellows enrolled in the first year of the fellowship programme. It received a 100% response rate. It asked fellows questions related to their use of information received by other fellows for use in advocacy efforts before and after becoming fellows, as well as about whether fellows had supported or collaborated on another fellow's advocacy or work or projects before and after becoming fellows. An increase in density values was significant for both information sharing and collaboration half way through the year-long program. By analyzing findings in relation to fellows' attributes, it was possible to identify whether fellows' use of information or collaboration was associated with their age, gender, geographical location or employment in government relative to academic or the non-profit sector, among others. The qualitative responses further validated these findings as fellows shared specific examples of use of information received from other fellows and of ways in which they had collaborated. In addition, participants indicated that the connections formed in this network were extended to projects outside of the fellowship programme. Fellows also provided some insights into what factors presented barriers to their networking within the Fellowship programme and these provide valuable insight into how the programme can be further improved for the next cohort of fellows.

Thursday, 4 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 068 Using Social Network Analysis for Evaluation Purposes**

**O 204 - Through Local Eyes: Integrating Social Network Analysis and fieldwork to reconstruct local configurations in evaluations**

*L. Tagle<sup>1</sup>, S. Celano<sup>2</sup>, S. Ghinoi<sup>3</sup>*

<sup>1</sup> *Italy's Presidency of the Council of Ministers, Evaluation Unit - Department for Cohesion Policy, Roma, Italy*

<sup>2</sup> *Independent Evaluator, Serafino Celano, Avellino, Italy*

<sup>3</sup> *Università di Bologna, Economics, Bologna, Italy*

European Cohesion Policy and European Rural Development Policy both emphasize the importance of an area-based approach, which rests on the responsibility and proactivity of local actors. Foundations, too, aim at strengthening local actors, building ties among them. Part of the evaluation of area-based interventions is based on understanding of relationships among local actors, describing local configurations, and capturing the ways in which they change over time. In particular, one of the goals of interventions is to strengthen the ties among organizations working in an area, in the idea that this may lead to both increased social cohesion and to better implementation of cohesion interventions.

Based on our experience in four evaluations in Italy's Mezzogiorno, the paper aims at discussing the issues involved in integrating Social Network Analysis (SNA) in fieldwork-based evaluation work. We needed to respond to local actors' knowledge needs in one of the four evaluations, while we performed the other three evaluations for an external donor. In all cases, the time-span we were interested in was rather long, ranging from 10 to 20 years. Also, we were working in conflict-ridden situations where actors offered widely differing interpretations of what had happened over time. We needed to assess the current state of the local social configuration, to identify the paths through which it had evolved over time, and to detect the contributions of interventions by public and private actors.

We adapted SNA techniques to fit in with the evaluation questions we faced and with the different situations in which we worked. We integrated it with predominantly qualitative analysis methods, to arrive at a mixed evaluation design. We could count on interdisciplinary teams. Finally, the results of SNA analyses and of the other strands of inquiry shed light on each other and produced further issues to be investigated, which, in turn, strengthened the findings. By presenting this paper, we are interested in sharing our experience and in eliciting a debate on how other evaluations have addressed similar issues.

Thursday, 4 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 069 Complexity and Systems Thinking

## O 205 - Toward more Resilient Project Evaluation: an Epic Battle between Professional and Political Logics

*O. Hassannejad<sup>1</sup>, L. Houghton<sup>1</sup>*

<sup>1</sup> Griffith University- Griffith Business School, Department of International Business and Asian Studies, Brisbane, Australia

This study seeks to shed some light on the increasing complexity of evaluation practices in the context of international development projects by exploring two competing and somehow contrasting professional and political evaluation logics.

This qualitative study is built upon an in-depth investigation of a phenomenon through the application of narrative inquiry approach with the conceptual aid of Actor-Network Theory (ANT) in the context of an international development project funded by the World Bank in Brazil. ANT as a sensitizing concept is used to examine the emerging and resilient role of different human and non-human actors (such as funder, sponsor, evaluators, government, project team, etc.) in the evaluation process of a large and complex international development project.

We found that evaluators and other key stakeholders deal with a constant battle between professional and political logics in which each actor pursues their own interests, attitudes and approaches through manipulation of evaluation product (report or text). Our study also shows how an emergent transformation of evaluation outcome occurs as a consequence of constant combat between professional and political logics. We articulate the evaluation practice as the collaborative sense-giving and sense-making processes in which both evaluators and evaluatees enhance their understanding of project and its outcome. Our study also proposes that the evaluation outcome could be either enlightenment, legitimisation, collusion or a combination of these three, depends on the result of battle between professional and political logics. We also integrate the whole findings into a concept map that demonstrates how different elements such as evaluation motivation, purpose, value, politics, process, temporality, criteria and actors give hand in hand to construct the evaluation logics.

We never sought to propose a generalizable concept model or framework in our study. Instead, we enjoyed experiencing a fruitful sense-making journey and are keen to share the joy with the evaluation community. We offer both project evaluation scholars and practitioners to go beyond scientific paradigms and objective-subjective duality and consider evaluation practice through the lens of social construction paradigm. Our study opens up a new avenue toward studying of evaluation logics in an emergent and complex environment and shows that the complexity of evaluation practice can be better understood through application of collaborative sense-giving and sense-making processes. However, future studies may complement or even contradict our findings through extending their research to other fields and context.

The collaborative sense-giving and sense-making concept map can assist evaluation practitioners particularly evaluators to enhance their understanding of the increasing complexity of evaluation practice. This study can also help the evaluation practitioners to adopt a more resilient and adaptive approach toward designing and implementation of evaluation practices to address the emergent nature of societies.

The originality of this study relies on transcending the conventional approaches toward project evaluation practice and enhancing our understanding of the two competing professional and political logics in the context of international development projects. The value of this research is on shedding some lights on the subjective process of translating evaluation product to outcome through the lens of sense-giving and sense-making.

Thursday, 4 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 069 Complexity and Systems Thinking

## O 206 - How to Evaluate Complex Adaptive Systems? Case: Disaster Preparedness and Disaster Risk Reduction

*P. Uusikylä*<sup>1</sup>

<sup>1</sup> *Frisky & Anjoy, Founder- CEO, Helsinki, Finland*

This paper introduces an alternative approach and methods to evaluate disaster preparedness and disaster risk reduction (DP/DRR) systems. Our evaluation carried out between 2016 and 2017 shows that by applying systems thinking, complexity theory and systems methods, we understand better the dynamics and interconnectedness of the DP/DRR. This applies both to interconnected risks (multi-risk landscapes) and interconnected actors (multi-actor networks). These systemic findings could thereafter be compared with formal program and project documents (logical frameworks and logic models) to see how well a priori expectations of the disaster preparedness actions matched with the “field reality”.

These results are part of the thematic evaluation commissioned by The Finnish Red Cross (FRC). The aim of the thematic evaluation was to promote institutional learning on successful/unsuccessful Disaster Preparedness and Disaster Risk Reduction (DP/DRR) project experiences and practices that can benefit better programming in the future. The overall objective of the study was to identify critical issues in designing, implementing and monitoring and evaluation by Finnish Red Cross and its partnering National Societies (NS).

This thematic evaluation consists of two main parts. First part presents the results of the meta-analysis and meta-evaluation of the ten countries and 17 projects. Meta-evaluation utilizes the IFRC/MFA evaluation criteria (relevance, impact, effectiveness, efficiency, sustainability and coherence). From this sample the final case studies were selected. Last part is the case study section introduced the findings and results of the field missions to the Philippines, El Salvador, Honduras and Tajikistan. Case study analyses use set of systems methods and tools to better understand the dynamics and interconnection between the risk factors and stakeholders in the field. These results will be presented in this paper. The systems approach utilized in case study section gave interesting insights of the dynamics and inter-connectedness of risk landscapes and inter-organizational DM networks. Study showed that by applying systems methods such as network analysis the risk components helped local disaster risk management units to better understand the interconnectedness of risk elements and the joint impact on those risks. Also, the relations and connections between the disaster risk agencies and stakeholders better explain why certain risk preparedness actions produce better results and effects. The more actors are connected to the network the more versatile the understanding of the risk preparedness and thus higher the resilience of preparedness actions.

Thursday, 4 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 069 Complexity and Systems Thinking

## O 207 - Systemic Evaluation in the Making: A Case Study

*M. Nieminen*<sup>1</sup>, *K. Hyytinen*<sup>2</sup>, *V. Salminen*<sup>3</sup>, *S. Ruutu*<sup>4</sup>

<sup>1</sup> VTT Technical Research Centre of Finland, Business innovation and foresight, Tampere, Finland

<sup>2</sup> VTT, Strategy, Espoo, Finland

<sup>3</sup> 4Front, Helsinki, Finland

<sup>4</sup> VTT, Business innovation and foresight, Espoo, Finland

Recently evaluation experts and scholars have paid increasing attention to systemic evaluation approaches to capture the increasing complexity of societal development (e.g. Forss et al., 2011; Patton 2011; Williams 2015; Mowles, 2014; Reynolds et al., 2016; Nieminen & Hyytinen 2015). The linear input-output-outcome -thinking included in the traditional evaluation does not correspond to the complex development processes and the multiple relationships between the contributing actors (Arnold, 2004). By paying attention to the interaction between various actors and to longitudinal complex dynamic behavior it is possible to explain how the complex interactions reduce, change or even hinder the desired outcome (Merrill et al. 2013).

In this paper, we suggest an approach combining three broad perspectives: futures view, systems view and multi-actor view. A futures view links evaluation with action, which is needed to realize desired future outcomes. A systems view helps to identify the dynamic interlinkages between different factors in the target of evaluation. Finally, a multi-actor view creates understanding of the forces that make desired action to happen, spread and gain foothold. (Hyytinen, 2017) In this context, a system is understood as a multi-dimensional entity following the ideas of so-called multi-level perspective of socio-technical change (e.g. Geels & Schot 2007; Geels 2002). Such dimensions and actors are e.g. politics and government, private enterprises, science, research and technology related institutions, and the third sector (NGOs) and citizens. The system is formed and maintained in a complex interaction between these actors and dimensions.

We have piloted this evaluation framework in Sitra, the Finnish Innovation Fund. Sitra works with wide-scale, future oriented societal development projects like a carbon-neutral circular economy and new working life and a sustainable economy. Sitra mostly facilitates social change by creating shared future visions, affecting politics and promoting interaction and networking among relevant actors in the system. Following from this the social impacts of Sitra are usually indirect and dependent on the actions and interaction of other social actors. Therefore, the evaluation focused on identification of impact pathways and systemic changes caused by Sitra. The evaluation was limited to one strategic objective in Sitra's activity: "Moving to a regenerative and collaborative economy". Sitra's role and contribution to the change was assessed in various system dimensions and how interaction between these dimensions was promoted to initiate a system level change. Special attention was paid to the multi-actor context of activity and stakeholder's views on Sitra's activity. To complement the analysis, a system dynamics model was developed to visualize the complex impact paths and interaction in the system.

The developed framework proved to be useful in making the underlying complexity of activities and their systemic interconnections visible and understandable. The framework: a) helped to understand the interlinked mechanisms inducing systemic impacts, b) made visible the wide spectrum of indirect impacts and their interconnections to wider societal impacts. The paper discusses the implementation of the framework, its challenges and advantages.

Thursday, 4 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 070 Promoting Evaluation Cultures and Ethics**

**O 208 - How to Become a Resolute and Assertive Evaluator:  
The Role of Ethos and Fundamental Texts in the Education  
of Evaluators**

*B. Neuhaus<sup>1</sup>*

<sup>1</sup> *evalux, Research & Development, Berlin, Germany*

Evaluators have to convince in various situations (e.g. during the evaluation process and in different fields and settings) and in different roles (e.g. as researcher, as facilitator, as service provider or as evaluator). They interact with many people in different functions and for various reasons (e.g. contract negotiation, data collection or counselling). They have to behave professionally and adequately – with everyone, at any time. This can be challenging especially when the object and questions of evaluation are highly demanding.

For this tough work evaluators need a solid knowledge, many competencies and skills. But in important situations they also need a strong posture and even ethos. Ethical codices, guidelines, principles and standards might contribute to its development. But do these fundamental texts really strengthen the evaluator? And do they back up evaluators when they are in a hustle? And how are they used in the education and training of evaluators?

In this contribution the author presents some insights and ideas to create the groundwork for an educational philosophy to help (emerging) evaluators not only to build up the necessary knowledge and to develop the necessary competencies and skills but also the awareness and posture in order to become resolute and assertive professionals who dare to stand up for good evaluation practice.

Thursday, 4 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 070 Promoting Evaluation Cultures and Ethics****O 209 - A Collaborative Implementation Evaluation among Governmental Instances in Brazil: An Innovative Experience for Health Care and Management Processes**

*G. Cardoso*<sup>1</sup>, *E. Moreira dos Santos*<sup>2</sup>, *E. Oliveira*<sup>3</sup>, *P. Passaro Toledo*<sup>4</sup>, *M.A. Santos*<sup>4</sup>, *D. Abreu*<sup>4</sup>, *C. Almeida*<sup>4</sup>, *S. César Luiz*<sup>4</sup>, *M. Alves*<sup>5</sup>, *J. Cardoso*<sup>5</sup>

<sup>1</sup> *Fundacao Oswaldo Cruz / ENSP, Evaluation Laboratory Laser, Rio De Janeiro, Brazil*

<sup>2</sup> *Fundacao Oswaldo Cruz / ENSP, Laser, Rio De Janeiro, Brazil*

<sup>3</sup> *Universidade Federal do Rio de Janeiro, Instituto de Estudos em Saúde Coletiva, Rio de Janeiro, Brazil*

<sup>4</sup> *Escola Nacional de Saúde Pública- Fiocruz, Laser, Rio de Janeiro, Brazil*

<sup>5</sup> *Ministry of Health, National Humanization Policy Coordination, Brasilia, Brazil*

Our work presents a collaborative evaluation experience – *AcolheSUS* Project – conducted between a Brazilian Ministry of Health (MoH) governmental instance, the National Humanization Policy (NHP), and the M&E Laboratory of the National School of Public Health (LASER), Fiocruz. The incorporation of M&E into government agencies is a challenge for evaluators and managers. Thus, the adoption of collaborative approaches is strategic in promoting an M&E institutional culture, assuming the inseparability between M&E capacity building and intervention (Cousins, et al., 2014, Santos Filho et al, 2009). The *AcolheSUS* Project aims to motivate the production of new management and health care practices, as to improve individuals' access to the Unified Health System (SUS), from three main entry points: Urgency and Emergency Network, Maternal Child Care Network and Psychosocial Care Network. Since the beginning, in 2017, the evaluators' team was invited to support the NHP/MoH management team in the implementation design of the intervention in 21 Brazilian states and develop a monitoring plan. Three pilot experiences were selected comprising the three networks (Urgency and Emergency, Maternal and Child and Psychosocial). Active pedagogical methodologies (Freire, 1987) were adopted at all stages, mainly through participatory thematic workshops, focused on the National Humanization Policy and its guidelines, as to define priority problems, participate in the work plan building and its subsequent modeling. This methodology was highlighted as an option for team qualification and conceptual alignment of the diverse groups of stakeholders involved. The three pilot experiences confronted us with different types of challenges. In the Psychosocial Care Network, the number, type of insertion and the variety of stakeholders was broader, including the participation of the patients of the Centre for Psychosocial Care. The patients' assembly was also included as an instance of discussion and resolution of the projects' issues. Shared modeling contemplated the richness and diversity of actions, actors and their networks. In relation to the Urgency and Emergency Network, the initially selected hospital was replaced in face of the internal difficulties encountered, such as the overlapping of projects not directly aligned with the *AcolheSUS* Project at the entrance door. The Maternal and Child Care Network included a state reference maternity hospital. There was strong adherence of several professional categories, mostly nurses, and the state management regional team. However, some actors were restricted, which made it sometimes difficult to align the work plan to the local health context. The preliminary results of the three pilot experiences indicate that the participation and integration of health managers, professionals and health system's users in the delimitation of local problems and the construction of a work plan not only facilitated a collective critical reflection, but produced innovative and resilient processes. Some changes have been observed in the routines of these health services, in the ambience of these places, and therefore, in the access of the individuals. Handling M&E concepts and understanding its importance in daily routine seems to facilitate their appropriation and use.

Thursday, 4 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 070 Promoting Evaluation Cultures and Ethics

## O 210 - Examination of Evaluation Culture

*C. Magdalina*<sup>1</sup>

<sup>1</sup> Ministry of European Funds, Programmes Evaluation Unit, Bucharest, Romania

**Rationale:** The presentation purpose is to share the experience gained by the Ministry of European Funds during a three year study on evaluation culture within the context of EU funded programmes implemented in Romania.

This study was triggered by the need to introduce a regular assessment mechanism for a technical assistance intervention aiming at developing an evaluation culture among the structures involved in managing EU funded development programmes. It was believed that a developed evaluation culture may have beneficial effects on the programmes' management.

**Objectives:** The presentation objective is to share our customised assessment methodology, including a set of dimensions, indicators, criteria and sub-criteria, which enabled the quantification of the Evaluation Culture Measurement Index (ECI) over three years, the identification of areas for further improvement and designing financial and non-financial interventions addressing the identified weaknesses. Not least ECI and its sub-components facilitated the identification and quantification of the impact of TA intervention of the progress achieved in the evaluation culture.

**Brief narrative and justification (with regards to the review criteria) [1]:** The concept of evaluation culture had to be clarified and structured around four dimensions – the demand side, the supply side, the level of dissemination and utilisation of evaluation results and the level of institutionalisation of the evaluation culture. Corresponding criteria (16), sub-criteria (30) and assessment indicators (64) were further developed and finally gathered in the so called Evaluation Culture Index. Minimum and maximum scores were established for each indicator and further, by aggregation, for sub-criteria, criteria and for the overall Index. Information sources feeding each indicator were identified using whenever possible the triangulation techniques. The TA intervention contribution to overall index was identified and quantified using documentary review, interviews and survey for each relevant indicator compounding the ECI.

The results of the study had many implications in the evaluation system, among which the review of the composition of the evaluation structures (evaluation steering groups), tenders selection criteria, new capacity building projects, including for data collection and statistics, are among the main outcomes.

The model can be transferred to assess the evaluation culture regarding national policies and development programmes. The full studies are available in Romanian and English, free of charge, on <http://www.evaluate-structurale.ro/en/k-tool/report-search>.

Thursday, 4 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

## S 071 Evaluating for Equity

**O 211 - Newborn and Child Care Along with Related Health-Seeking in Bihar: An Exploration by Gender of the Child***T. Mahapatra<sup>1</sup>, A. Raj<sup>1</sup>, A. Das<sup>1</sup>, G. Sai Mala<sup>1</sup>*<sup>1</sup> CARE India Solutions for Sustainable Development, Concurrent Measurement and Learning, Patna, India

**Rationale:** Position of women in the family and society are considered critical while addressing community behaviors aiming at a positive deviance in the newborn and child health related indicators. Evaluation methods in public health sector, owing to its generalizability are capable of generating representative pictures regarding health practices, health-seeking and variations thereof across gender-strata to explore any gender-specific predilections in those practices

**Methods:** Community-level newborn and child health related frontline-worker-driven (FLW) counselling and related practices were explored through four annual rounds of a serial cross-sectional household survey, during 2014–17 involving a statewide representative sample (N>78000) of women who delivered during last two years stratified by five separate age groups of the children (0–2, 3–5, 6–8, 9–11 and 12–23 months). Stratum-specific distribution and associational analysis using regression modelling were conducted to determine variations of the relevant service coverage and practices across strata of the gender of the last-born child.

**Findings:** Women who delivered a male baby during last three months were more likely to receive postnatal FLW counselling for dry cord care [Adjusted Odds Ratio (AOR)=1.11 for 0–2 in 2017, p<0.0001], delayed bathing (AOR=1.16 for 0–2 in 2017, p<0.0001), exclusive breast feeding (AOR=1.13 for 0–2 in 2017, p<0.0001), timely initiation of complementary feeding (AOR=1.11 for 6–8 in 2017, p<0.0001), age appropriate diversity of complementary feeding (AOR=1.07 for 6–8 in 2017, p<0.0001) as opposed to their counterparts who delivered a girl during the same period. Interestingly, while their practices regarding delayed bathing (AOR=1.07 for 0–2 in 2017, p<0.0001) were also showing the similar pattern, it was reverse for exclusive breast feeding (AOR=0.84 for 0–2 in 2017, p<0.0001) and continued breast-feeding (AOR=0.91 for 9–11 in 2017, p<0.0001). Further-more bottle-feeding (AOR=1.22 for 0–2 in 2017, p<0.0001; AOR=1.23 for 3–5 in 2017; p<0.0001, AOR=1.26 for 0–2 in 2017, p<0.0001 and AOR=1.15 for 0–2 in 2017, p<0.0001), formula/animal milk-feeding (AOR=1.24 for 0–2 in 2017, p<0.0001), cereal-based /formula-food (AOR=1.44 for 6–8 in 2017, p<0.0001, AOR=1.70 for 9–11 in 2017, p<0.0001) and milk-products (AOR=1.25 for 6–8 in 2017, p<0.0001, AOR=1.20 for 9–11 in 2017, p<0.0001) were also found to be more common for male babies compared to girls in the same age group. With reference to females, male babies were found to suffer more from diarrhea (AOR=1.02 for 0–2 in 2017, p<0.0001) and acute respiratory infections (AOR=1.14 for 0–2 in 2017, p<0.0001, AOR=1.20 for 3–5 in 2017, p<0.0001). In terms of health-seeking, care-seeking from unqualified practitioners was the preferred options for sick female children, while for sick males qualified private providers (AOR=1.59 for 3–5 in 2017, p<0.0001, AOR=1.37 for 9–11 in 2017, p<0.0001) were often preferred. Boys were also more likely to receive full immunization (AOR=1.08 for 12–23 in 2017, p<0.0001) as opposed to girls.

**Conclusion:** Newborn and child care related FLW counselling and relevant practices were found to vary across gender strata. Male babies appear to receive preferential care, especially where potentially higher financial resources are required.

Thursday, 4 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 071 Evaluating for Equity**

**O 212 - Dealing with Flexibility and Complexity: Lessons from Impact Assessment of the Shaurya Dal Programme of Government of Madhya Pradesh, India**

*M. Singh*<sup>1</sup>

<sup>1</sup> UN Women, Monitoring and Evaluation, New Delhi, India

Shaurya Dal was formed by the State of Madhya Pradesh in India, in response to the increasing incidences of crimes against women. Madhya Pradesh is the second largest state in India. According to the 2011 census, the tribal population in Madhya Pradesh was 73.34 million, constituting 21.1% of the total population, making women and girls from the marginalised communities even more vulnerable. An analysis of the situation indicated that violence against girls and women can be curbed only by securing administrative, social and community partnerships. Shaurya Dal was launched in 6 pilot districts on 24<sup>th</sup> June 2013. In year 2015–16, the programme was up-scaled to all the 51 districts of Madhya Pradesh. Shaurya Dals are community-based groups, which respond to and prevent Violence Against Women and Girls (VAWG) in their community

UN Women in 2017 conducted an Impact assessment of the programme. The assessment focused on processes of formation and mobilisation of Shaurya Dal members and their role in identifying, informing and addressing the incidence of VAWG, while also improving women's/girls' awareness of and access to entitlements and benefits. It further analysed Shaurya Dals' approach to improve the responsiveness of the community to ending VAWG, while also influencing attitudes and beliefs among families, communities and society for the creation of a safe environment for women and girls. In all, it generated evidence for the Government for further development of policies/strategies to strengthen existing interventions and/or advocacy initiatives on ending VAWG.

The learnings from the assessment will give the audiences an insight into the complexities associated with applying evaluation methodologies to assess State supported community based programmes working on gender equality and empowerment of women. The socio-economic landscape of Madhya Pradesh coupled with the existing social practices that perpetuate gender stereotypes and inequality, made the assessment of the impact of the programme particularly challenging. The complexity of the intended and the unintended issues that the programme seeks to address, combined with the change in mindsets and behaviours as a result of the programme, makes it difficult to use a structured approach to assess the impact. The government programmes in India are often marred by the lack of baseline data and thorough monitoring systems. In such a case the job of an evaluator becomes even more difficult, the assessment gives an insight into how to deal with the complexities while being flexible in approach.

The assessment focused on factors that lead to violence against women that operate at individual, relational, community and societal levels. The findings suggest the need for interventions operating at these various levels. In conclusion, assessment also bring to forefront the issue that VAW not only affects women and girls, but makes the society vulnerable to both the direct financial costs of dealing with the phenomenon and the indirect productivity costs that result from it, therefore ending VAWG is a key component to create a resilient society that achieves its full potential.

[i] [http://www.mp\\_gov\\_in\\_sc\\_st\\_welfare\\_2001/tribal/](http://www.mp_gov_in_sc_st_welfare_2001/tribal/)

Thursday, 4 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 072 Evaluating Longterm, Systemic Change**

**O 214 - Evaluating Impact in Complex Environment – Dialogue on the Needs, Methods and Practices Based on Sitra's Impact Evaluations**

O. Oosi<sup>1</sup>, K. Vataja<sup>2</sup>, P. Parkkonen<sup>2</sup>, S. Rausmaa<sup>3</sup>

<sup>1</sup> Owl Group Ltd., Evaluation, Helsinki, Finland

<sup>2</sup> Sitra, Strategy, Helsinki, Finland

<sup>3</sup> Owl Group Ltd., Evaluation, Helsinki, Finland

This paper presents a dialogue between the evaluation commissioner and the independent external evaluators on the needs, methods and practices of evaluating impact on systemic changes. The systemic societal change as a starting point of evaluation poses challenges to the traditional evaluation practices. The paper rethinks the design and methodological issues of analysing the impact of actors, which have set long-term goals for their societal impact.

Traditional forms of the evaluation fail due to complexity nature of systemic changes. Systemic changes require lot of dedicated actions across the different sectors which over time build upon each other, and they involve multiple partners and stakeholders with changing roles. To analyse impact on systemic changes we need methods that consider the dynamics of the operating environment and the time span of societal changes which is usually long.

Our case is based on wide scale evaluations of Sitra, which is a Finnish future-oriented think-and-do-tank. Sitra's operations are guided by six principles of sustainable well-being, all of which are interlinked and form the shared goals for impact. The evaluation is targeted towards these goals. Our examples will present why and in what situations the traditional approaches and project and program-based designs are not sufficient.

We suggest that the evaluation design benefits from the approaches such as the evaluation of strategies or that of principles focused evaluation. In the case presented, the evaluators have used mixed-method approaches for tracing and harvesting the contribution of the evaluand using data from various sources.

Thursday, 4 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 072 Evaluating Longterm, Systemic Change

## O 215 - Using a Wide Lens to Take a Long View: How Systems Thinking Can Help Incorporate Long-Term Perspectives into Short-Term Evaluations

*A. Koleros<sup>1</sup>*

<sup>1</sup> *Palladium, Research- Monitoring and Evaluation, London, United Kingdom*

Taking a long-term perspective on change can offer a number of benefits for programme managers and evaluators, such as showing multiple and interconnected causal patterns and in detecting impacts which only emerge over time. For many evaluators, however the pressing question is not “does a long-term perspective add value”, but rather “how to take a long-term perspective in practice”?

In recent years there has been a growing number of funders adopting long term evaluations, with funders such as the UK Department for International Development (DFID) demonstrating an increased interest in long-term evaluation exercises, particularly in looking retrospectively at the impact of development cooperation initiatives over a long time horizon (Kolers, et al, 2016). Although this is an encouraging trend, donor funding cycles and other organisational factors continue to play a large role in how evaluations are commissioned – and for how long. The reality is that long-term evaluations will likely continue to be the exception rather than the rule in the near term.

However, commitment to a long-term perspective in evaluation does not imply that commissioning a long term evaluation is the only way to “take the long view” in practice. This session contends that taking a systems lens can help a short term evaluation integrate a more long term perspective.

First, using the case of an evaluation of a police reform programme implemented in the DRC and a subsequent sustainability study conducted two years after programme completion, we will describe how an actor-based system perspective within a short-term evaluation can assist in assessing the likelihood of long-term programme sustainability.

Next, we explore how the development of multiple and overlapping system maps can be used to more comprehensibly identify potential intended and unintended long-term impact pathways beyond a short-term intervention period, drawing from experiences evaluating the impact of development interventions on long term economic and social change processes over a 40-year period in the eastern region of Nepal.

We conclude that building these long-term systems perspectives into short-term evaluation can contribute to evaluations which better assess the likelihood of programme sustainability, more robustly estimate contributions to long term change processes, and more reliably identify unintended consequences and alternative explanations.

We argue that this type of analysis is both feasible and necessary for more credible short and long-term program evaluation. We hope that these experiences contribute to an improved consideration of incorporating systems perspectives in both short- and long-term evaluations.

Thursday, 4 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 072 Evaluating Longterm, Systemic Change**

**O 216 - Evaluating the Sustainability of Complex Interventions:  
The Case of Police Reform in DRC**

*D. Stein<sup>1</sup>, A. Koleros<sup>1</sup>*

<sup>1</sup> *Palladium, Research- Monitoring and Evaluation, London, United Kingdom*

This paper examines the tension inherent in evaluating the short term versus long-term effects of programme interventions. Short-term impacts can often emerge as a result of direct delivery to beneficiary populations, while the effects of system-level interventions are often slower to materialise. These challenges are particularly relevant in the context of increasing demands from donors for development programmes to demonstrate results as well as a growing recognition among evaluators of the need to engage with 'complex' change processes. Understanding the sustainability of these changes is key to the ability of evaluation to full assess the extent to which an intervention improved the resilience of a society. Yet, this aspect is often overlooked in evaluation methods that adopt more standard forms of mixed methods data collection within a programme's implementation period.

This session explores this tension through the evaluation of a police reform programme implemented in DRC in 2016 and subsequent sustainability study conducted two years after programme completion. These tensions are common when considering security sector reform processes, which often include a combination of supply side system-level interventions as well as direct community level demand-side activities. By combining surveys and longitudinal, qualitative outcome monitoring, the initial evaluation design sought to capture the programme's complexity within an overall theory-based design while estimating the specific impact of the interventions. From this basis, the session will then illustrate how the findings gathered through these methods formed the framework for assessing the programme's sustainability using a political economy analysis.

The session will then detail the conclusions drawn from these combination of methods; namely that the programme achieved population-level change in the short term, including improvements in the public perceptions of safety and security among women and marginalised groups in particular. However, the programme's failure to account for the dynamics and incentives facing individuals in institutions responsible for driving reform reduced the sustainability of these achievements. The session will conclude with reflections on the benefits and challenges of employing this combination of evaluation methods when assessing both security and justice interventions as well as programmes designed to build social resilience more broadly.

Thursday, 4 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 073 Turbulent Contexts and Evaluation 2. Adaptations in Evaluation and Programming**

**O 217 - Managing Evaluation of Complex Interventions:  
Resilience in Evaluation**

*S. Farina<sup>1</sup>, C. Lovato<sup>2</sup>*

<sup>1</sup> *Broadleaf Consulting, Evaluation, Vancouver, Canada*

<sup>2</sup> *The University of British Columbia, School of Population and Public Health, Vancouver, Canada*

**Rationale:** In today's world of complex interventions evaluators are often called upon to plan, implement and report on multiple, sometimes staggered, programs using different approaches in multiple sites, and all while coordinating different levels of participation of a diverse range of stakeholders. This complexity has resulted in a need for evaluators to expand their scope to manage these projects effectively. This presentation describes key challenges and approaches to building resilience into the management of evaluation of complex interventions. It will include two case examples and a discussion of lessons learned.

**Objectives sought:**

- Discuss challenges and approaches to complex evaluation using examples from a variety of multi-program, multi-stakeholder evaluations.
- Explore approaches to managing evaluation of complex interventions and supporting institutional strengthening and capacity building.

**Brief narrative and justification:** Conducting evaluations in turbulent times is stretching evaluator roles in a variety of ways, including managing complexity and fostering resilience. The literature on project management suggests research in this area has shifted from project control to project adaptability. This presentation explores key challenges and approaches to managing evaluation of complex social interventions that include multiple programs and sites, multiple stakeholders, multiple indicators, different implementation timelines, regional variations, and a range of stakeholders. Adaptive approaches to evaluation such as Principle Based Evaluation and Most Significant Change technique will be explored as a means to support resilience, good judgement, learning, and critical thinking in organizations. The presenters will share two case examples involving complex evaluations. We will discuss lessons learned including capacity building within client organizations and creating feedback and learning loops.

Thursday, 4 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 073 Turbulent Contexts and Evaluation 2. Adaptations in Evaluation and Programming****O 218 - Evaluative Resilience Through OODA Looping***C. Elkins<sup>1</sup>*<sup>1</sup> *Belling the Cat- Ilc, International Evaluation, Chapel Hill, USA*

**Rationale:** Resilience is constructive adaptation to cope with or overcome challenges, shocks, and their consequences. Some responses to such turbulence merely add to the chaos, or provide a thin illusion of innovation. More useful responses, however, recalibrate our perspectives and priorities with tools that increase operational resilience in the face of volatility. Robust evaluation design is resilient to normal patterns of change, but not necessarily to fluctuations in more volatile environments. This paper presents key benefits to evaluation that are potential in the synergies between evaluative strategies and an approach designed to equip people to make effective decisions, demonstrating operational resilience, even under extreme uncertainty.

**Objectives sought:**

- Clarify key aspects of resilience as an adaptive strategy
- Present the Observe, Orient, Decide, Act (OODA) looping system
- Articulate OODA as a model of resilience particularly useful for ambiguous/precarious contexts
- Highlight benefits for contemporary evaluation of explicitly integrating OODA models

**Brief narrative and justification:** Rapidly escalating interpenetration of international development assistance and conflict or peace-precarious situations in the early years of the 21<sup>st</sup> century dovetailed with increasing interest in complexity and chaos models to conceptualize the contemporary environments in which we perform evaluations. While a great deal of creative thinking has focused on these areas, it can be difficult for the average evaluator to glean the useful tools and functional approaches they can adapt and practice for specific evaluation studies. There continues to be a dearth of practical yet rigorous approaches for sound design and quality implementation of evaluations in the vast majority of cases which are not catastrophically chaotic but nonetheless require adapting successfully to intrinsic uncertainty and unanticipated inconstancy in circumstances.

John Boyd created a sophisticated strategic learning system that is sometimes oversimplified to four steps abbreviated as OODA. As one way to approach decision-making under uncertainty, OODA explicitly integrates constructive ways to think about many challenges familiar to those working in the contemporary evaluation context, such as strategic opposition, imperfect knowledge, and change over time. It is a toolkit that, with practice, builds capacity to narrow uncertainties and respond nimbly to new information – whether challenges or opportunities.

The OODA mindset, in other words, embeds resilience and embraces adaptation within a strategic context. For evaluators and other evaluation stakeholders, this reorientation keeps the focus on results, and measurable progress toward them, but explicitly models the dynamic and interactive nature of real-world operations. Numerous benefits accrue, including reducing the disconnect between idealized evaluation designs and the actual evaluation that takes place; increasing opportunities to learn and share learning from innovative processes and unanticipated outcomes; and reprioritization of learning impact as a core evaluation value.

Thursday, 4 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 073 Turbulent Contexts and Evaluation 2. Adaptations in Evaluation and Programming**

**O 219 - Evaluating Complex Systems for Strengthened Resilience: Experiences from the Africa Risk Capacity**

*C. Simon*<sup>1</sup>, *J. McConnell*<sup>2</sup>, *M. Tarazona*<sup>3</sup>, *F. Le Quesne*<sup>3</sup>

<sup>1</sup> *Independent, Monitoring and Evaluation, Boulder- CO, USA*

<sup>2</sup> *Independent, Monitoring and Evaluation, Kansas City, USA*

<sup>3</sup> *Oxford Policy Management, Disaster Risk Management, Oxford, United Kingdom*

This paper discusses the design and implementation of a formative evaluation of the African Risk Capacity (ARC), an innovative, sovereign-level disaster-risk insurance mechanism. The evaluation is the first of several DFID-commissioned formative and impact assessments planned over ten years. The paper explores issues of evaluation related to *programmatic, implementation environment, and evaluation utilisation* complexity, all of which pose challenges to the design of a robust and flexible evaluation approach. The learnings from this evaluation provide insights into (i) how methodologically to evaluate resilience-building initiatives occurring in complex environments, and (ii) how evaluation can be used to promote resilience by strengthening the processes needed to improve country and community responses to disaster.

Established in 2012, ARC uses index-based weather risk insurance products combined with technical programming to help member states improve their capacities to better plan, prepare, and respond to extreme weather events and natural disasters. Ultimately, ARC aims to help vulnerable households be more resilient to disasters through the receipt of timely, government support. Seven countries have joined one or more risk pools and four countries have received an insurance payout. ARC's *programmatic complexities* are driven by high levels of interdependence across multiple stakeholder levels (household, community, country, region) and types (government, NGOs, donors, private sector), and across a diverse and interrelated set of processes, policies, and systems. ARC's *implementation environment complexities* include uncertainty around how and when disasters manifest in each member country. Uncertainty also arises in that ARC's impact is conditioned both by the decisions of multiple, interacting agents and by the relative strength of processes and systems operating outside of ARC's immediate control. For example, the nature and strength of in-country programme delivery mechanisms to households in response to disaster vary widely. Finally, *evaluation utilisation complexities* emerge when considering the role of the evaluator. There is need to balance serving as an embedded partner in the learning and adaptation objectives of the formative evaluation with maintaining objectivity and scope for the longer-term impact evaluation.

Framed by a set of evaluation questions, the evaluation uses contribution analysis as a means of testing the links in the theory of change. Data collection was organised around three work streams- an organisational review, country case studies, and a global perspective assessment. In addition to discussing the methodological components and challenges of the evaluation, the authors conclude by reflecting on the importance of evaluation not only in determining programme effectiveness; but in understanding and evaluating the underlying processes needed to build resilience. Findings suggest that there are tangible barriers to change within countries related to political will, governance, and a lack of understanding and therefore mistrust of insurance products, all of which must be overcome for resilience to take hold. ARC must find new ways to address the short-term incentives of political decision-makers which are often mismatched with the longer-term value proposition of insurance. This formative evaluation typifies evaluation for learning and sets an example of how to establish the groundwork for collaborative knowledge-sharing and adaptive programming of a complex programme.

Thursday, 4 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 074 The Political, Cultural, and Commercial Aspects of Evaluation

## O 220 - The Evaluation of Public Policies and the Politics of Evaluations: Views from Postmodernist and Complex Realities in Latin America

*S. Salinas*<sup>1</sup>

<sup>1</sup> RELAC, Executive Committee, La Paz, Plurinational State of Bolivia

This presentation addresses a different dimension of the evaluation of public policies: The politics of evaluation. Some of the related issues and questions refer to the implicit and oftentimes disputed interests that lead decisions and investments in evaluations, the logic, approach and methods selected, the underlying assumptions, the public or private nature of the evaluation process and results, the expected use...or non-use of the findings. Furthermore, how does this all relate to accountability, transparency, participation and inclusion in a specific context, and how this in turn relates to historical backgrounds, citizenship and governance systems and models.

To address the politics of evaluations it is important to recognize that public policies represent a specific response to public affairs in particular contexts, and that these express decisions that relate to how public issues are identified and interpreted, and about the option considered best to address them. Public policies are thus social constructs that express world views and decisions of (subjective) actors under specific situations and power-relations. As there are no neutral, objective actors that define public policies, there is no such thing as neutral or objective public policies...Furthermore, there is no such thing as objective, "non-political" evaluations of public policies. Evaluations reflect world views, power relations and political interests. Recognizing the political nature of evaluations can enhance their legitimacy.

Thus, the evaluation of public policies implies addressing its nature and assumptions, the process of its generation and its capacity to embrace today's citizens in increasingly multi-cultural realities. This implies new conceptual, methodological and operational frameworks for an integrated understanding and participative processes that expand the sense of belonging and ownership. Citizens no longer see themselves as "beneficiaries" of public policies but as active actors and, as such, also active evaluators. This goes beyond the traditional rationale of participative processes, to embrace key governance, equity and stability issues under a new social paradigm.

In Latin American dynamic, multi-cultural contexts, urbanization, migration, modernization, inclusion and globalization processes have resulted in new or recently visibilized & recognized postmodern and multi-identity actors with increasing expectations: rural & urban indigenous groups, multiplied gender & sexual orientations, migrants, youth cultures, global netizens...are just some examples of the diverse and complex map of today's citizens that demand society's recognition and effective public policy response to their (new) needs and aspirations. This complexity implies new paradigms and a new framework of social relations, which not only taint the definition of public policies, but their evaluation: the objectives, the rationale, the questions, the methods and the assumptions. It is not only an issue of generating new policies for new actors, it is also about new lenses to address emerging and re-signified actors, and about new perspectives to address traditional issues.

Thus, in the context of complexity and diversity, this presentation aims to contribute to the discussions and proposals around transformative, rights based evaluations and the SDG mandate of "leave no one behind".

Thursday, 4 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 074 The Political, Cultural, and Commercial Aspects of Evaluation**

**O 221 - Cultural Competency and Responsiveness in Evaluations: Promoting Resilience and Development in Challenging Contexts**

*H. Paulmer*<sup>1</sup>

<sup>1</sup> Goss Gilroy Inc, GGI, Ottawa, Canada

As evaluators, we play key roles in promoting resilience and furthering social change. Evaluations cannot be culture free. Cultural competence in evaluation is necessary and important for evaluators of all backgrounds. Cultural differences exist as we evaluate programs in diverse and challenging international contexts. Evaluators, working in contexts different from their own cultures, are continually challenged to explore not only their practices but also their beliefs and assumptions about evaluation. Cultural incompatibility can contribute to lack of trust, and decreased engagement of stakeholders and beneficiaries.

Without an appropriately situated understanding of cultural factors and how they potentially influence evaluation results, an evaluator may find it challenging to accept culturally responsive evaluation theory and approaches. Understanding cultural context is necessary for accurate interpretation. All aspects of evaluation process take place in cultural contexts. Evaluative understanding and judgments have to be grounded in culture and context. Evaluating with validity requires cultural competence. Failure to address culture threatens validity and thereby the ability to promote resilience and development in cross-cultural and challenging contexts.

Culture is at the core of evaluation process. Embracing multiple cultural perspectives in designing/conducting evaluations is essential. Communities that are diverse ethnically and culturally make key contributions to evaluations. The lens we (evaluators) are using, influences analysis and reporting. Do one's values, beliefs, assumptions, and practices shape the evaluations? How does one adopt/adapt in challenging cross-cultural contexts? As evaluators, how do we deal with flexibility and complexity? In addition to touching upon concepts, frameworks, cross-cultural aspects, the paper presents lessons learned from extensive experience working across various diverse cultures in four continents, evaluating projects/programs in challenging and development contexts. The paper also highlights "essentials" to work effectively across cultures and being culturally responsive in evaluating initiatives in cross-cultural and challenging contexts.

Culture is a relevant aspect irrespective of the evaluation framework. Cultural competence in evaluation theory and practice is critical for the profession and for the greater good of society. As evaluators, cultural competence is crucial to gaining the trust of communities in collecting data and during the entire evaluation process, which enhances our ability to report in a culturally responsive manner. The presentation also addresses issues of diversity, equity, and equality aspects in cross-cultural and challenging contexts. The paper highlights the importance of being relevant and responsive as evaluators.

**Selected Bibliography:**

Hood, S., Hopson, R.K., & Frierson, H.T eds. (2015) *Continuing the Journey to Reposition Culture and Cultural Context in Evaluation Theory and Practice*. InfoAge Publishing Co.

American Evaluation Association (2011). *Public statement on Cultural Competence in Evaluation* by American Evaluation Association.

Kirkhart, K.E (2011, May). *Missing the mark: Rethinking validity threats in evaluation practice*. Paper presented at the annual meeting of the Eastern Evaluation Research Society, Absecon, NJ.

Symonette, H. (2004). "Walking Pathways toward becoming a Culturally Competent Evaluator: Boundaries, Borderlands and Border Crossings." *New Directions for Evaluation* 1002:95–110.

Thursday, 4 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 074 The Political, Cultural, and Commercial Aspects of Evaluation

## O 222 - Skin in The Game: Epistemological, Political and Moral Contradictions In International Development Evaluation Discourse

*J.M. Statman*<sup>1</sup>, *A.E. Ansell*<sup>2</sup>

<sup>1</sup> Northern Nigerian Education Initiative, Education, Abuja, Nigeria

<sup>2</sup> Emerson College, Institute for Liberal Arts and Interdisciplinary Studies, Boston, USA

Discourse is a path from one contradiction to another: if it gives rise to those that can be seen, it is because it obeys that which it hides. To analyze discourse is to hide and reveal contradictions... (M. Foucault, *The Archeology of Knowledge*, 1972, p. 151)

**Rationale:** In recent years, monitoring, evaluation and learning have become universally accepted as essential components of international development interventions including those sponsored by major donors such as USAID (US), DFID (UK), European Union, GIZ (Germany), SIDA (Sweden) and the World Bank and many tens of millions of Euros, Pounds and Dollars of development assistance resources aimed at assisting and empowering impoverished, marginalized and vulnerable populations are consumed annually on these endeavors. As the necessity of evidenced-based “M&E” has now assumed the status of the virtually unquestionable, unsurprisingly, there have been few attempts to critically and contextually reflect on the discourse, ideology and practice of international development evaluation, its implicit assumptions, epistemology, ethics and value.

**Objective:** The focus of this paper is a consideration of what is to be seen and what becomes hidden in the discursive enterprise of evaluation. It addresses the dynamics, dilemmas and contradictions inherent in the conduct of evaluative acts, focusing on the conduct of “monitoring, evaluation and learning” as performed in the arena of international development activities implemented in fraught and contested contexts characterized by widespread poverty and marginalization, endemic conflict and corruption, massive inequality, political oppression and resource exploitation.

**Organization:** The paper commences with an analysis of the power dynamics and consequences inherent in the fundamental positioning of evaluation as judgmental “other”: observing, monitoring, measuring, appraising, advising, and the range of epistemologies and resultant roles and practices evaluators have adopted and performed as conditioned by the evaluation profession’s ideological, institutional, financial and careerist habitus, within the specific context of poverty, oppression and powerlessness.

It then describes the international development enterprise itself: in terms of interventionist practice (sectoral projects and programs) and institutional/financial frameworks contextualized within the overarching dynamics of market-driven globalization.

The paper concludes with a critical analysis of the role, narrative position and professional and financial stake and career of the evaluator/evaluation construct as discursively constructed and conceptualized, performed, and embodied by the discourse of “evidenced-based” results-focused, “value-for-money” international development practice. Drawing on examples from their work in countries including Nigeria, South Africa, Zimbabwe, Guyana, Albania, Egypt, Tunisia and Azerbaijan, the authors identify conceptual, political, ethical and operational dilemmas and contradictions implicit in the discourse of international development evaluation and consider possible alternative constructs.

Thursday, 4 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 078 Role of Evaluation in Governance, Accountability and Innovation**

**O 223 - Small but Mighty? The Market for Evaluation in the Canadian Not-For-Profit Sector**

*C. Elliott<sup>1</sup>, S. Heath<sup>2</sup>, L. Robert<sup>3</sup>*

<sup>1</sup> University of Ottawa, Telfer School of Management, Ottawa, Canada

<sup>2</sup> University of Ottawa, Faculty of Education, Ottawa, Canada

<sup>3</sup> REL Solutions Inc., Consultant, Ottawa, Canada

Canada's Not-For-Profit (NFP) sector is a critical segment of Canada's evaluation industry. This sector is quite diverse, comprised of four groups: i) private foundations, ii) public foundations, iii) charitable organizations, and iv) non-profit organizations. It represents an estimated 170,000 organizations in Canada, half of which (54%) are run by volunteers, with the remainder employing approximately 2 million people. The majority of these organizations, however, are very small, with the top 1% of organizations commanding 60% of all revenues (Imagine Canada and Philanthropic Foundations Canada, 2014). Due to the diversity of this sector, it is very difficult to get a national "pulse" on the supply and the demand for evaluation in this segment.

The objectives of this paper are to discuss the findings of a qualitative research study that aimed to: (a) explore the nature and extent of evaluation work conducted by NFP organizations in Canada, (b) describe the market forces and factors that affect evaluative inquiry in this sector, (c) highlight the degree to which evaluative work in this sector influences the overall market for evaluation market in Canada, and (c) identify research priorities to further investigate NFP evaluation in Canada. To guide our initial inquiry, we identified the following broad research question: To what extent does Canada's NFP sector engage in evaluation? To answer this question, we followed a multi-stage methodological framework. As part of a larger study on the evaluation industry in Canada, we reviewed the published and grey literature on NFP evaluation and conducted qualitative interviews with key experts working in the field.

The findings of this study suggest that the sector may be "small but mighty". That is, the NFP sector constitutes a small segment of the overall Canadian market for evaluation services and there appears to be relatively little systematic program evaluation being performed. Primarily driven by accountability pressures from donors, NFP's have struggled with limited capacity and chronic underfunding. Yet, despite these challenges, NFP's have demonstrated a keen interest in program improvement, understanding 'what works', evaluation capacity building, and implementing innovative approaches to evaluative inquiry and performance measurement (e.g., developmental evaluation, collective impact). These findings illustrate the need to further explore evaluation activities in the NFP and the factors that drive such activities, such as availability of resources and prioritization of evaluation over performance monitoring activities. It would also be instructive to tease out differentiating factors between each of four sub-sectors, (private and public foundations, charitable organizations and non-profits), to examine the unique characteristics of each.

This type of research will equip those funding evaluation activities and those responsible for conducting evaluation activities in the NFP sector with data needed to spark and encourage change and advancement in the field. To ensure that this paper is of interest to individuals working in various sectors, we have also deliberated on the factors and drivers that influence evaluative work in general.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 075 Evaluating Investment Facilities

## O 224 - Evaluation of Financial Sustainability of Investment Facilities

*V. Coppens*<sup>1</sup>

<sup>1</sup> ADE s.a., Head of Evaluation Department, Louvain la Neuve, Belgium

**Rationale:** Investment facilities (IFs), i.e. regionally or thematically focused financial instruments that support projects promoting the development of the private sector and commercially-run public enterprises, have increasingly become key modalities for development partners in the context of SDGs and Agenda 2030. They indeed contribute to achieving a sustainable and inclusive economic growth, which is a key aspect of building resilient societies.

Their financial revolvability is a challenge, especially when investing in higher risk operations or countries, like in Low Income Countries and countries in situation of fragility. Evaluating their financial sustainability is a challenge too.

**Objectives:** This presentation aims to share ADE's approach and instruments developed for assessing financial revolvability of Investment Facilities.

**Rationale and justification:** Assessing financial sustainability is a challenge. Specific mechanisms or instruments are not always used by investment facility managers.

ADE has developed specific approaches and instruments for two IF evaluations, notably one for assessing the financial sustainability of the EIB-managed ACP IF and one for the financial revolvability of the Infrastructure Development Fund (IDF) managed by FMO.

» To assess the financial viability of the EIB ACP IF, a specific approach was used, including

- A mapping of the distribution of the risk for the signed exposures of the IF (cumulative distribution per loan, from acceptable to equity type), and
- An analysis of the income statements of the IF over the evaluation period.

The financial sustainability assessment revealed that:

- A priori all operations were designed so as to fully recover the financial costs (except the concessional element of the interest subsidies);
- There was little management of the risk at portfolio level;
- There was evidence of due attention to the constraint of financial viability and revolving character of the IF;
- In some cases, this character was given more weight than the objective of maximising the use of funds for development purposes.

» ADE has built a revolvability model for IDF. It addressed two requirements:

- to provide an objective and realistic view of how IDF 2 (from 2019) sustainability/revolvability might be achieved, and
- to be used by FMO as a way of managing IDF 2 should it go ahead.

The model has been built on a set of key assumptions (incl. on priority sectors, annual commitments and impairment rates). This evaluation is being finalised at the time of drafting this abstract (March 2018).

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 075 Evaluating Investment Facilities

## O 225 - Great Expectations: An Enquiry into the Premise of Directing Institutional Investors Towards Emerging Markets, Especially in Times of Market Crises

*R. Narayanan*<sup>1</sup>

<sup>1</sup> World Bank Group, Senior Evaluation Officer, Washington, USA

**Background:** To move the needle on emerging market development, private sector capital needs to be mobilized in billions through new investment vehicles with appropriate risk management and regulatory frameworks (Source: United Nations Financing for Development Conference, Addis Ababa, 2015). Global financial assets exceed \$300 trillion with nearly \$100 trillion invested with sovereign wealth funds and institutional investors such as pension funds, insurance companies and fund-of-funds (Source: Institute of International Finance, 2017). One of the most intriguing areas of development financing today is to find unique ways to channel the aforementioned trillions of commercial, investable capital to developmental projects in emerging markets. In some instances, such long-term capital can potentially double up as an “accelerant” for a multilateral bank (MDB)’s or international financial institution (IFI)’s direct interventions to respond to a global or regional crisis. A recent strategy, pursued by MDB/IFIs in this context, is to position emerging market private equity both as an investable asset class that can generate positive returns for stakeholders and a contributor to development goals.

**Objectives:** The aim of this paper / presentation is **three-fold**: (a) to present methods and techniques to evaluate and benchmark the performance of funds, (b) to examine the potential of private equity funds as a crisis-response tool and (c) to share specific examples where such methods and instruments have been used to evaluate business lines within the World Bank Group.

**Motivation:** The presenter has many years of career experience in private sector and related evaluations, having recently pioneered a new benchmarking methodology to evaluate collective investment vehicles and investment funds. At the UNDP-NEC Conference in Istanbul in 2017, Raghavan conducted a pre-conference workshop on private sector evaluation along with Fredrik Korfker which got a lot of attention by the conference participants, and was recognized by the UNDP organizers as one of the top 2 workshops in the conference based on participant survey.

**Audience:** The audience for the pre-conference workshop on private sector evaluation will be evaluation professionals with experience with the private sector or who intend to concentrate more on private sector development. In the context of the Sustainable Development Goals the private sector will be very important in creating resilience environments and markets. Therefore, learning more about the current methodologies, applied at the moment in the leading private sector development finance institutions, and being able to take part in in-depth discussions on the challenges of private sector evaluation can enrich the workshop participants.

**Presentation outline (45 minutes-60 minutes)** The presentation will set out of the global context of institutional investors, the role of private equity instruments as a vehicle, its application in emerging markets context, describe the various approaches to engage with private sector through the above instruments, detail the various methodologies used to analyze the interventions’ effects, with respect to the market crises environment and provide suggestions in a forward looking way.

**Teaching methods:** Presentations will be given, and it is expected that there will be a lot of interaction between the presenters and the participants.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 075 Evaluating Investment Facilities**

**O 226 - Method Study and Methods Engineering Techniques  
in Improving Business Processes and Outcome Driven  
Organizational Viability**

*S. Premakanthan*<sup>1</sup>

<sup>1</sup> *Symbiotic International Consulting Services SICS, Consulting & Professional Development, Ottawa- Ontario, Canada*

One of the ways a business or an organization could contribute to the bottom line is to manage its key business processes in a cost-effective manner. Business processes make things happen when executed in the most economical and efficient way. They are the heart and soul of a business as they lead to the production of goods and services which benefits its customers or clients. **(Rationale)**

It is an introduction to the application of a very powerful process improvement technique known as Method Study or Methods Engineering. They are used very commonly by Work Study Practitioners and Industrial Engineers all over the world. Method study is a major branch of Work Study. The other is Work Measurement, also known as Time and Motion Study. In the 1990's Method Study was broadened in scope and referred to as Business Process Re-Engineering. **(Justification: Rethinking evaluation methods, design, and criteria and Combining methods in evaluation)**

As a Professional, I have applied the techniques & tools to many industrial and white collar work processing situations to improve productivity. Outcome Management professionals engaged in purpose driven policies, programs, projects and initiatives could benefit from these techniques in their Measurement, Monitoring & Evaluation practice. **(Objective)**

**Note to Reviewers:**

**The abstract was accepted to be presented at the EES 2016 Conference. Presenter was unable to attend due to personal reasons and informed the organizers in time.**

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 075 Evaluating Investment Facilities****O 227 - Impact Evaluation of Rural Finance Institution Building Programme on Food Security Status Among Beneficiaries in Anambra State, Nigeria**K. Chah<sup>1</sup><sup>1</sup> *University of Nigeria Nsukka, Department of Agricultural Extension-, Nsukka, Cameroon*

Rural Finance Institution Building Programme (RUFIN) which became effective as a pilot study in 2010 in 12 states of the Federal Republic of Nigeria, is currently in its impact period. The programme is a strategic means by which the rural micro financing sector will be developed and strengthened in order to deliver adequate, efficient and sustainable financial services to the rural poor which may go a long way to improve food security. Information on food security of households in Nigeria is continuously needed for early famine warning purposes so as to plan and target interventions appropriately. This need has become more critical with the prolonged *Boko Haram* insurgency in North East Nigeria and the widespread economic recession across the country. This makes it important to evaluate the impact of RUFIN on the food security status of beneficiaries household in Anambra State; ascertain perceived constraints to beneficiaries in utilizing RUFIN services; and identify possible strategies for improving performance of RUFIN. Significant difference between food security status of beneficiaries (RB) and non-beneficiaries (NRB) was determined. Multistage, purposive and snowball sampling techniques were used to select 120 households (60 beneficiaries and 60 non beneficiaries) for the study. Data were collected using structured interview schedule. Also, key informant interview and personal observations were employed. To determine the food security status of the beneficiaries and non-beneficiaries, food security survey module revised in 2012 was used. Descriptive statistics, t-test and factor analysis were used to analysis data. About 43% of RB and 22% of NRB were food secured. A significant ( $p < 0.05$ ) difference was observed between RB and NRB in the nature of food eaten in the household ( $t = 1.37$ ), while no significant ( $p > 0.05$ ) difference was observed between the two groups in the number of times households feed in a day ( $t = 0.00$ ) with food insecurity score of  $t = -1.53$ . Loan term, managerial and system embedded, were identified as constraining factors to full utilization of RUFIN services by beneficiaries. High interest rate, embezzlement of funds by group top officials and gender centeredness of RUFIN services were the major variables that loaded under each factor for loan terms, managerial and system embedded factors, respectively. Giving farmers loan at subsidized rate and strengthening the legal backings of group activities were suggested by beneficiaries as possible strategies to enhance utilization of RUFIN services. The food security status of the beneficiaries is better than that of the non beneficiaries, therefore the programme should be lunched in other states of the federation taking into consideration the major constraints and suggested strategies to enhance the utilization of RUFIN services. The result of the study has provided necessary information on the programme impact with which policy makers, donor agencies, and intervention programme planners can use as working tool in initiating, sponsoring and executing better policy for the development of mankind.

Thursday, 4 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 076 Innovative Methodologies and Techniques for Causal Understanding**

**O 229 - Combining Realist Evaluation and Causal Loop Diagramming in Evaluating a Performance-Based Financing Intervention in the Ugandan Health Care Sector**

*D. Renmans<sup>1</sup>, N. Holvoet<sup>1</sup>, B. Criels<sup>2</sup>*

<sup>1</sup> *University of Antwerp, Institute of Development Policy IOB, Antwerp, Belgium*

<sup>2</sup> *Institute of Tropical Medicine, Department of Public Health, Antwerp, Belgium*

**Rationale:** Performance based financing (PBF) in the health care sector of low-and middle-income countries (LMIC) is a supply-side reform package that consists of performance-based financial incentives for health providers (facilities and/or workers) combined with a separation of functions (purchaser, provider, verifier), (spending) autonomy for the health facilities, strict monitoring and verification of services, community involvement, result-based planning and accountability arrangements. Indeed, it is a multi-component intervention, which, moreover, highly interacts with the context in which it is being implemented. We therefore opted for a combination of realist evaluation and causal loop diagramming in order to grasp both the different mechanisms that are being triggered by the intervention components and the interaction and systemic effects of these interactions.

**Objectives:** Firstly, we evaluated the BTC/Enabel PBF intervention in Western Uganda and looked for what works for whom, when, where and how. A second objective was to show that causal loop diagramming can be instrumental in helping evaluators to answer the 'how'-question in a realist evaluation. A final objective was to look for shortcomings of the methodological strategy and come up with solutions for them.

**Methods:** We gathered data before the start of the intervention and after one year of implementation. We used quantitative surveys and semi-structured interviews with the health workers, key informant interviews and observations. We analysed the data according to the theorized program theory and mechanisms, which were constructed based on an extensive literature review, the program manual and key informant interviews at the level of the Ministry of Health, the donor, and the governing authorities of the private not-for-profit facilities. Both the program theory and the mechanisms were visualized and partly analysed using causal loop diagramming.

**Results:** Our study found that four mechanisms were particularly relevant in the explanation of observed outcomes: the accreditation mechanism, the 'saliency' mechanism, the financial accessibility mechanism, and the management mechanism. Factors related to the pre-existing health care system, the design of the intervention, and the capacities of the stakeholders were important barriers to the triggering of some of mechanisms in some of the facilities. The use of the causal loop diagrams showed to be useful in analyzing the interactions between the mechanisms but also to visualize and describe much more clearly a complex program theory and context-mechanism-outcome configurations.

**Justification:** The evaluation of PBF in LMIC will only become more important as new and more thorough evidence is being asked for by both supporters and opponents. Similarly, realist evaluation and systems thinking in the health care system of LMIC are just in their infancy and certainly its combination has only been used by a few evaluators. It is thus safe to say that this was an innovative approach that has not been performed in this way before, but which has great potential for future evaluations. We also found some limitations to the approach and propose pathways to overcome them. In the future we aim to refine and further strengthen the methodological strategy and we hope that the conference can be an interesting starting point.

Thursday, 4 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 076 Innovative Methodologies and Techniques for Causal Understanding**

**O 230 - Assessing the Pertinence of Primary Healthcare Programmes' Objectives using a Hierarchical Fuzzy Approach: a case study in Brazil**

*A. Jatoba<sup>1</sup>, L.Z. Oliveira<sup>2</sup>*

<sup>1</sup> *Fundação Oswaldo Cruz, Centro de Estudos Estratégicos, Rio de Janeiro, Brazil*

<sup>2</sup> *Fundação Oswaldo Cruz, Presidência, Rio de Janeiro, Brazil*

Evaluation in complex systems like healthcare must consider a set of interdependent variables and indicators, as well as their in feedback loops, which affect the decision environment and bring new information and consequently new decisions.

In addition, when there are multiple evaluators and multiple criteria available, conflicts of opinion are common, as each evaluator has his or her own opinion about the problem and the importance of each criterion to assess the evaluation question.

Thus, the Fuzzy Logic can provide a way of representing the human decision mechanism in a more realistic way, allowing subjective evaluations, approximations and inaccurate values, expressed in natural language. The Fuzzy Logic describes logical systems by placing the values selected by the evaluators in subsets of a given interval. Thus, evaluators can express their opinion verbally using common linguistic terms, and mathematical models, presenting adequate results to rules that combine evaluators' opinions, can process such values.

In summary, fuzzy numbers represent verbal expressions, i.e., values that people express in natural language. This representation is useful for characterizing complex problems, in which the evaluation is rough or within an inaccurate scale. Therefore, the Fuzzy Logic is suitable for the evaluation of pertinence, because it enables the positioning of the element within a context represented by a range of values, instead of requiring the choice of a single value

We used the hierarchical Fuzzy model suggested in this study in a case study in a Brazilian primary care program called "More Doctors". The final objective of the More Doctors Program is to guarantee the quality of access to Primary Healthcare services, universally.

In the case presented in this study, the evaluation of the pertinence of the strategic objectives of the More Doctors Program uses a set of quality indicators. These indicators are imprecise and subjective, which hinders logical inferences. Thus, Fuzzy Logic showed promising in the construction of a model that allowed the aggregation of the participants' opinions.

Data for this study is collected through the application of a structured questionnaire to a group of 10 specialists. Participants' answers fed the Fuzzy model, which aggregates the answers considering the weights of the evaluators and the relative weights of each of the criteria, indicating the relevance of the strategic objectives of the program with values that range from "Not pertinent" to "Very pertinent".

Thursday, 4 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 077 Evaluation Capacity Building 2**

**O 231 - The Link between Evaluation Policy and Organisational Capacity for Evaluation: What Evaluation Scholars and Practitioners are Saying**

*H. Al Hudib<sup>1</sup>, B. Cousins<sup>1</sup>*

<sup>1</sup> *University of Ottawa, Faculty of Education, Ottawa, Canada*

During the current period of public fiscal restraint growing competitiveness in the global economy, and increasing conditions of unpredictability and complexity, the level of international interest in evaluation capacity building (ECB) continues to grow among governments, civil society organisations, programme and policy funders (demand side) and evaluation practitioners (supply side). Evaluators and evaluation scholars who are committed to building and integrating evaluation culture and practices in organisations have been adding to the growing base of theoretical and empirical ECB literature. Many studies have theorized that the enhancement of organisational evaluation capacity is essential to the efficient production of quality evaluations that enable organisations to meet their increasing accountability requirements and to use evaluation for organisational development and decision making. Although a considerable progress has been made in understanding how to ensure ECB is effective and in identifying its role, there is a lack of empirical knowledge about the role played by evaluation policy in the organisational capacity for evaluation. Practitioners and scholars of evaluation, therefore, have recently called attention to the central importance of evaluation policy and its impact on most aspects of evaluation practice and theory (Trochim, 2009). Nevertheless, the relevant literature shows that evaluation policy is not yet an integral part of discourse in evaluation literature on how to improve evaluation practice or enhance organisational capacity for evaluation.

The purpose of this paper is to address this gap by drawing attention to evaluation policy and its influence on organisational capacity for evaluation through empirical inquiry. In this paper, we present research findings based on a set of 18 interviews with leading evaluation scholars and practitioners from Europe, Canada, and the United States, who have made substantial contributions to what we know about ECB and/or evaluation policy. We asked about evaluation policy's role in the integration of evaluation into the organisational culture, and contextual facilitators and barriers that influence its role. The findings of this study are critical for developing the basis for a more focused and in-depth understanding of evaluation policy and its role in leveraging organisational capacity, and for expanding our knowledge about the facilitators and barriers that influence the role of evaluation policy in enhancing the integration of evaluation into organisational culture. Given that ECB has an explicit goal of developing organisational capacity to do and use evaluation, it is critical that we expand and deepen our knowledge about the process by situating it within a broader context of evaluation policy.

Thursday, 4 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

## S 077 Evaluation Capacity Building 2

### O 232 - The Role of Evaluation for Supporting the Most Vulnerable in Kosovo

*R. Kahlert<sup>1</sup>, S. Danaž<sup>2</sup>*

<sup>1</sup> European Centre for Social Welfare Policy and Research, Health and Care, Vienna, Austria

<sup>2</sup> European Centre for Social Welfare Policy and Research, Welfare and Work, Vienna, Austria

The role of evaluation is incompletely developed in Kosovo, a fledgling democracy that still struggles under the conditions of partisanship, favouritism and intransparency. The young state is also characterized by the context of decentralization of public responsibilities, but of limited accompanied funding. The European Union is dedicated to financing projects for fostering a more resilient Kosovar society, including a project on how to improve the social net and to better social services for the most vulnerable groups in Kosovo: the elderly, the disabled, children, and victims of violence (again, especially women and children). In this particular project, an external organization (the European Centre for Social Welfare and Research, based in Austria) was tasked to perform research and evaluation at various stages and levels of project planning and implementation. This presentation reports on the project's evaluative tasks as well as the future need to foster local capacity in evaluation.

The initial evaluative task was an assessment to determine the status quo of the current needs regarding social services. The evaluation team involved stakeholders at the central level, municipal Centres for Social Work and NGOs providing social services. The team used statistical data as well as interviews and questionnaire responses to determine existing and perceived needs for social service delivery. Next, the team established the unit cost of social services and designed a financial formula for distributing the necessary funds to the municipalities. While these two steps were not clear-cut evaluations, they entailed evaluative components, e.g., assessing whether a current vulnerable group had received the right amount of services, and if not, how this needed to change.

Evaluation, public policy, and social work are disciplines that come into play when aspiring to make the Kosovar society more resilient. It is therefore crucial that evaluators are part of this process and not just watching by the sidelines. In this case, the international project team has the necessary, although not explicitly required expertise in evaluation. Furthermore, it engaged local social science experts who embraced and performed the evaluative tasks.

In emerging democracies such as Kosovo, evaluation's relevance lies in its special role of establishing and guaranteeing the rule of the law. At first, this role is often played by external organizations such as by our team. In the long run, however, evaluation capacity and skills need to be developed within the particular country. There is currently no national formal cross-disciplinary Volunteer Organization for Professional Evaluation (VOPE) in Kosovo. In the future, the conditions for building a strong professional network of evaluators and evaluation stakeholders need fostering. We outline how a project like ours fosters local capacity building in evaluation by encouraging local experts in establishing professional evaluation structures, through the application of evaluation tools to exemplary problems, and by emphasizing the importance of evaluation for democratic resilience.

Thursday, 4 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 078 Role of Evaluation in Governance, Accountability and Innovation**

**O 233 - The Role of Innovation in Supporting a Resilient Healthcare System in the UK; Can Evaluation Help (and, if So, How)?**

*T. Ling<sup>1</sup>, S. Marjanovic<sup>2</sup>*

<sup>1</sup> RAND Europe, Head of Evaluation, Cambridge, United Kingdom

<sup>2</sup> RAND Europe and THIS Institute, Innovation Health and Science, Cambridge, United Kingdom

Evaluating the 'big ticket' issues of the day is important if resilience is to be strengthened. However, it inevitably brings challenges of its own sort that EES members need to debate. This paper will offer an opportunity to take this debate forward.

The authors spent three years working on a major study/evaluation of the healthcare innovation landscape in the UK. In this paper they report on the overall conclusions but primarily reflect on how best to ensure that evaluative activities can be oriented towards meeting the needs of decision-makers in an environment where policies shift during the lifetime of the evaluation, the organisational setting change, the knowledge base develops and the concerns of policy-makers evolve. These are almost inevitable consequences of delivering large scale evaluations on topics of considerable political importance in relatively turbulent times and the authors reflect on how evaluators/researchers can still generate evaluations that add value.

Starting from the end-point of a piece of work that was largely relevant, helpful and robust, the authors are interested in how governance and research processes supported learning and adaptation within the project and what more could have been done to support this. It also considers if and to what extent these processes created a centre of gravity for the study which was as useful as possible or whether, conversely, some potentially important questions were insufficiently explored. Finally, the authors will make recommendations for how future research and evaluations of this sort might be funded and supported. However, rather than resolving these important questions, our hope is that this paper opens up further debate on this important topic.

Thursday, 4 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 078 Role of Evaluation in Governance, Accountability and Innovation**

**O 234 - Post NPM: The Evaluation of Trust-Building Governance**

*V. Denvall*<sup>1</sup>

<sup>1</sup> *Linnaeus University, Social Work, Vaxjo, Sweden*

Trust-based and collaborative governance is suggested to replace NPM and the audit-society. This kind of governance is expected to increase the responsibility and influence of employees as well as end-users, reducing control and audit.

The rationale of this paper is to discuss evaluation in times of post-NPM. The paper is based on reviewing evaluations of trust-based governance in health care and social work in Sweden. A committee is framing trust-based governance and a dozen evaluations have been conducted. The paper specifically relates to the conference's theme on design and management of evaluation in turbulent times. Evaluation of trust-based governance in this new governance-landscape needs to take into account at least three intertwined aspects; globalization, organization and innovation.

Globalization and digitizing contribute to increasing social complexity and require the ability of transformation within established organizations. Collaboration between actors and users is expected to result in innovative services. This is assumed to be in contrast to traditional administration ensuring legalism and accountability. Instead, risk-taking and creative attempts are emphasized in order to involve actors who not normally are represented. According to this view, new forms of organizational cultures need to be built that support collaboration and co-creation. Network management and multi-level management complement or replace bureaucracy. New skills and knowledge should be brought into innovation systems creating new social values. Co-production emphasizes cooperation with services-users when organizations seek to increase their absorptive capacity.

Turbulent times imply the need for innovation to change working methods, organization or forms of cooperation. Innovations are linked to trust, since the transformation of knowledge into new goods and services may occur when actors meet in processes characterized by reciprocity and respect. Broad interdisciplinary environments and different kinds of alliances will produce knowledge alongside traditional scientific research. A more local knowledge production will then emerge. The new services that are created are expected to break with traditional patterns and, therefore, partnerships are being organized with participants from several spheres and levels: private, public and users.

How then does evaluation correspond to the world of post-NPM? Traditional outcome evaluations might be improper. A hypothesis is that evaluators are creative combining multiple methods that favor dialog and mutual understanding. I will analyze the methods and theories that have been in use by the evaluators. Specifically I will discuss the use of criteria in the evaluation of trust-based governance and how results have been used. Domains and logics together with changing conditions and cognitive systems within and between organizations create barriers but also new opportunities for evaluation.

Thursday, 4 October 2018  
11:30–13:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 079 Professionalization and Evaluator Competencies****O 235 - A Case Study of an Evaluation Association's Pathway to Professionalization: AEA's Past, Present, and Future***S. Tucker<sup>1</sup>, J. King<sup>2</sup>, D. Podems<sup>3</sup>*<sup>1</sup> AEA Board member- AEA Competency Task Force- and President- Evaluation & Development Associates LLC, Evaluation, Pasadena, USA<sup>2</sup> University of Minnesota, Organizational Leadership- Policy- and Development, Minneapolis, USA<sup>3</sup> OtherWISE, Research and Evaluation, Cape Town, South Africa

**Rationale:** Voluntary Organizations for Professional Evaluation (VOPEs) outside of the US have taken the lead on developing the professional status of program evaluators. Global practice suggests professionalization efforts by VOPEs can be conceptualized as an iterative than linear sequence of building capacity in three blocks: 1) Evaluation norms, standards, and guiding principles, including reference to evaluator or evaluation team requirements; 2) Framework of competencies, capabilities, and qualifications for individual evaluators; and 3) Defining processes to verify whether individuals meet the required competencies, capabilities, and qualifications. Ultimately, professional status is established through these three blocks and should culminate in formal recognition (Davies & Brümmer 2015) and practice that is contextually adaptive and appropriate (Picciotto 2011).

The third block has proven to be challenging for many VOPEs. While each case is contextually unique, VOPEs actively navigating block 3 can help other VOPEs be better prepared to formally recognize and support their members competencies. Countries such as New Zealand, South Africa, Japan and Canada illustrate different pathways to formally recognize individual evaluator competencies as does EES's regional effort with Voluntary Peer Review (VEPR) of evaluators.

**Objective:** The objective of this paper is to present a case study of how a VOPE is navigating entry into the third block. The case in question is the American Evaluation Association (AEA). This work is relevant to the AEA membership because, following approval by the AEA Board, it is the members who will vote in 2018 to approve the set of competencies. The looming question is how will these competencies be used? It is relevant to the field of other VOPEs to learn from our successes and be forewarned about challenges to anticipate in promoting resilience and action at member, association and public good levels.

**Case Method:** In response to the growing international discussion of evaluator professionalism, AEA's Board created the Competencies Task Force (CTF) in 2015 to: 1) develop competencies building on AEA guiding principles and standards as well as 2) begin a thoughtful discussion amongst AEA membership about appropriate next steps in the professionalization of program evaluators, a challenging task given the diverse practice of AEA's 7400 members, one fifth who are international. Over the past three years, AEA's CTF has reviewed existing sets of national, regional and global competencies for program, policy, and personnel evaluation to identify and frame foundational competencies. The Task Force revised the competencies three times using systematic engagement protocols and surveys with members, affiliate organizations, and other VOPEs.

This paper will describe the past three years of navigating the first two building blocks, validating our framework and mobilizing for the third building block. Member and stakeholder engagement protocols were used to validate competencies, explore how the competencies can be used, identify membership concerns, seek insights from other VOPEs, and analyze other VOPE formal recognition systems (e.g., methods of recognition, their use by individual evaluators, commissioners and other stakeholders, and efforts to build the validity, credibility and manageability of competency initiatives). Recommendations about next steps will be sought from the EES audience.

Thursday, 4 October 2018  
11:30–13:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 079 Professionalization and Evaluator Competencies**

**O 236 - Reinforcing the Link Between Learning and Action:  
Engaging Students in Grounded Learning About Evaluator  
Competencies and Evaluation Practice**

*J. Lavelle*<sup>1</sup>

<sup>1</sup> *University of Minnesota-Twin Cities, Organizational Leadership- Policy- & Development: Evaluation Studies, Minneapolis, USA*

The field of evaluation has grown in recent years, as evidenced by an increase in the number of professional evaluation organizations worldwide and greater demand for evaluation services in a variety of contexts. In tandem, much empirical work has been published in recent years on evaluator competencies (e.g., King et al, 2001, LaVelle, 2014, Stevahn et al, 2006, Stevahn et al, in progress), and evaluator competencies are an important topic in many evaluation-specific courses (Davies & LaVelle, in preparation). Less clear, however, is how to make evaluator roles and competencies immediately relevant for students in university.

The essential competencies for evaluation practice explicitly or implicitly undergird the courses and programs designed to prepare future evaluators for scholarship and practice (King et al, 2006a; LaVelle, 2014). Most professors of evaluation would likely agree that the overt integration of essential competencies is important, though it can be a challenge to help students see the through line between what is taught in the classroom and how the different pieces are useful in evaluation practice. Simultaneously, teachers of evaluation are also juggling the usual academic challenges with classroom management alongside the requirements for tenure.

The root of this study combines the need for practical classroom management techniques (e.g., taking attendance), empowering students with job search strategies, and illustrating the link between what's taught in the classroom and what is sought in the employment marketplace. The author will describe 1) the theoretical rationale for requiring classroom participation to further the students' learning experiences, 2) how requiring students to conduct job searches lead to cognitive and attitudinal change, 3) how faculty can require students to submit job ads as a proxy for attendance, and 4) how student-led analysis of the collected job ads at the end of the semester can be used to reinforce learning and provide an ongoing needs assessment for the current educational curriculum. Implications for the evaluation profession, practitioners, and evaluation education programs will be discussed.

Thursday, 4 October 2018  
11:30–13:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 079 Professionalization and Evaluator Competencies****O 237 - Strengthening Evaluative Thinking Skills for Achieving the Sdgs in the “Post-Truth Era”**C.A. Asenjo Ruiz<sup>1</sup><sup>1</sup> *United Nations Office on Drugs and Crime, Independent Evaluation Unit, Vienna, Austria*

The presentation aims at stressing the importance of developing evaluative thinking skills within evaluation capacity building efforts for enhancing the contribution of evaluation to SDG achievement while mitigating the negative impact of “post-truth Era” principles on evidence-based decision-making. In order to illustrate this idea, UNODC’s new eLearning course on evaluation –which pays special attention to building evaluative thinking skills– will be briefly described. In recent years it has been pointed out the importance, especially in the context of the 2030 Agenda for Sustainable Development, of grounding policy making, management, and evaluation processes on solid evaluative thinking in order to foster the demand of evaluations, increase the ownership and quality of evaluations, and ensure that evaluation results actually feed into decision-making processes. The threat to evidence-based decision-making posed by “post-truth Era” principles – according to which objective facts, under certain circumstances, are less influential in shaping public opinion than appeals to emotion and personal belief– makes even more apparent the need to mainstream evaluative thinking capacity building in national and international efforts for strengthening national evaluation capacities. UNODC’s Independent Evaluation Unit has the mandate to contribute to improved accountability and evaluation-based decision making in the response of UNODC and Member States to illicit drugs, crime, and terrorism – domains where using critical thinking skills is paramount. Building upon its extensive experience in evaluating such programmes, the UNODC is developing a new e-learning course that aims at enhancing evaluation knowledge and skills of UNODC staff and government stakeholders. The course addresses both the need to improve the understanding of how to conduct and use evaluations as well as the need to enhance evaluative thinking skills of key stakeholders throughout the evaluation process. Conceived as a self-paced, online tool, the course offers trainees the opportunity to acquire (or refresh) their knowledge of basic evaluation concepts along with an overview of common thinking biases that should be avoided in order to make the most of evaluations. The training also aims at contributing to strengthened national evaluation capacity building at the country level for achieving the SDGs, as mandated by the United Nations GA Resolution A/RES/69/237.

Thursday, 4 October 2018  
11:30–13:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 080 Applying the DAC Criteria in Evaluation

## O 238 - Development Evaluation Beyond Aid

*V. Borges*<sup>1</sup>

<sup>1</sup> *Fundação para o Desenvolvimento e Intercâmbios Internacionais, President, Praia, Cabo Verde*

**Rationale:** There is an ongoing debate and a collective sense that OECD-DAC evaluation criteria should be revisited. Most of the issues raised are technicalities aimed to:

- Update the criteria to a fast-changing world and ensuing political and development challenges (growing political uncertainty and social complexity, 2030 agenda, etc.).
- Better frame these criteria to the reality and conditions in the field. Beside these, there are other issues that have not been brought up in the discussions.

**Objectives:** This paper will examine some of these issues articulated around three objectives: i. Discuss the expression development evaluation. ii. Analyse the perspective behind DAC evaluation criteria and practices. iii. Outline the notion of “endogenous development evaluation” and subjacent additional criteria.

**Brief narrative and justification:** Most of time, the literature speaks about development evaluation to refer to project/programmes evaluation funded by donors. Development goes beyond aid. What about country funded programmes? How to deal with donor's and country's perspectives in a creative and fecund way to promote a better understanding and mastering of development process? The development landscape in developing countries is composed of various actors: foreign, national, governmental, institutional, private and non-governmental. It is then worthwhile asking whose standpoint prevails in evaluation. Traditionally evaluation initiatives and design are linked to plans, projects and programmes promoted/funded by donors according to their rules, procedures, terms of reference and experts. In spite of unquestionable efforts to integrate recipient country institutions, experts and civil society actors, empirical observation points out that donor's perspective is mostly dominant. Does this mean that foreign perspective is misplaced or ill-conceived? Or are they just incomplete? If yes, what is missing? Developing countries cannot have development ownership without endogenous evaluation! Why is country-owned evaluation marginal or even missing in most developing countries? Is there lack of political, institutional and social awareness concerning evaluation? What are the prerequisite for endogenous evaluation? Without denying or downplaying the importance of DAC criteria for country-owned evaluation, an endogenous approach will perhaps bring about new interpretation of these criteria or additional ones that will make evaluation more comprehensive. In this case, which additional criteria will/should be added? The paper will try to elaborate and make the case for endogenous evaluation – both institutional and independent. There is urgency to go beyond the donor's perspective to respond to the imperative to strengthening national capacities for development evaluation for both governmental and non-governmental actors. I believe endogenous evaluation will contribute to a better understanding of development and the role of national and foreign actors in development management and planning. It has the potential to reshape the dialogue between donors and recipient countries making it more authentic and fruitful. The paper will also include suggestions to donors and developing countries on how to better promote and mainstream endogenous evaluation. This, I believe, will raise social and political awareness, concur to better use of domestic resources, strengthen government transparency and accountability and speed up the implementation of Paris Declaration on aid effectiveness (2005).

Thursday, 4 October 2018  
11:30–13:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 080 Applying the DAC Criteria in Evaluation**

**O 239 - Stretching the Limits – Challenges and New Options  
in Applying the DAC Criteria to the Evaluation of Human Rights  
Interventions**

*M. Raab<sup>1</sup>, J. Rocha<sup>2</sup>*

<sup>1</sup> Camino gGmbH and Michaela Raab consulting, International evaluation, Berlin, Germany

<sup>2</sup> Camino gGmbH, International Evaluation, Berlin, Germany

Terms of reference for evaluations in the international development sector tend to focus on the OECD/DAC evaluation criteria (relevance, sustainability, effectiveness, efficiency, impact and sustainability) and their definitions as enshrined in DAC publications. Most of the time, evaluation TOR in the sector list all DAC criteria; often, long and heterogeneous lists of evaluation questions come with each criterion. Frequently, the timing of the evaluation, its time frame and available resources do not match the expectations reflected in those lists of evaluation questions. Furthermore, the resulting lack of focus often affects the use of evaluation findings for program learning negatively.

While widely accepted and used criteria certainly are useful, the DAC criteria have limitations when it comes to interventions in the political and social realms, such as work to promote human rights or gender equality. We argue that there is a need to critically rethink how the DAC criteria can be used and how they could be amended to better support the design of evaluations that trigger program learning for more resilient communities.

Interventions designed to further human rights come with complex pathways and patterns of change that include reversals and backlashes, but also unexpected opportunities. Overarching goals in work on human rights and gender justice tend to be located far beyond the reach of a single project or programme; even achieved outcomes may vanish when their context changes. In these domains, applying the DAC criteria is neither straightforward nor necessarily useful; in the opposite, the uncritical use of DAC criteria and their definitions can be misleading.

Our paper will illustrate these challenges with examples from our recent experience with a large international portfolio evaluation in the area of human rights and gender justice programming. We will explain how we dealt with those challenges in that evaluation and other examples from our practice. We will discuss how the DAC criteria might be used more flexibly and where adaptations are needed to create evaluation processes that support joint learning in turbulent times. For instance, we will show how the efficiency criterion can be addressed meaningfully in human rights and gender equality interventions with fluid costs and uncertain benefits.

Thursday, 4 October 2018  
11:30–13:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 081 Innovative Approaches to Data Collection

## O 240 - Mobile Data Collection is the Future for Evaluation but Often in Remote Poor Areas it is Often Fraught with Problems

*S. Nyamhuno*<sup>1</sup>

<sup>1</sup> USAID Southern Africa, Evaluation and learning, Johannesburg, South Africa

The following lessons were gleaned from a data collection that was conducted in a small indigenous community of Jouberton, in South Africa in February 2017 using mobile devices.

#### A. Size of gadgets

If the field team is going to be using the fingers, it is important that they get bigger gadgets because smaller phones are not compatible with those with fat fingers. They might end up choosing the wrong options in drop down menu. Gadgets with pens are a better option.

#### B. Training

It is important that the field team is used is trained to be comfortable with the gadgets before embarking on actual data collection process lest they take a lot of time in one place.

#### C. Safety

It is important to secure the gadgets against theft. It would be good to have gadgets that are of no street value or that cannot be used for any other purposes so as to reduce pilferage risk especially in very risky places.

#### D. Network

Mobile data collection requires reliable network or else the evaluation team will not be able to access the app. It is important to ask gatekeepers about a reliable network before settling for one.

#### E. Data security

To minimize the loss of data there is need to have the data saved on the server as soon as the interview is over. Better still, it's better to have a system that captures live onto the SQL server. As a back-up, it is imperative to have an off-line mode; should the network fail, one can always synchronize when the network is back.

F. Testing -The gadgets and the questionnaires need to be tested on a separate population other than the actual target population. This gives the evaluation team a sense of how the actual exercise will pen out.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 133 Handling Complexity: Approaches and Methods

## O 242 - Evaluation Methods for the R&D Programs Dedicated to Solve Social Problems

*S.C. Byeon*<sup>1</sup>

<sup>1</sup> KISTEP Korea Institute of Science and Technology Evaluation and Planning, Office of S&T Policy and Planning, Seoul, Republic of Korea

This research is on the evaluation method for the R&D program dedicated to solve social problems. This program is to provide practical goods or services to solve social problems related with everyday life in the area of social welfare, living environment, or public safety. Especially this program provides not only proper technologies but also total solution including improvement in laws or regulations, development of services and the delivery system. With these kinds of R&D programs, Korean government is trying to enhance social contribution of R&D by changing the purpose of R&D from economic growth to happiness of people and quality of life. Therefore different evaluation method and indicators were applied to evaluate the feasibility of this R&D program and suggest future direction.

This program was initiated on 2014 with total budget US\$ 42 million for 4 years, and later the government tried to continue the program with total budget US\$ 90 million for another 5 years. So, in this research, the performance of the program was reviewed and feasibility of the future R&D plan was investigated if it is necessary and proper for national R&D.

To review the performance of the program, the achievement of the program goal, the fulfillment of the previous recommendations, research management system and environmental factors were investigated. Especially cooperation between related government departments, efficiency and effectiveness of the program, and similarity and overlap with other similar programs were investigated in detail.

To investigate the feasibility of the future plan, technological, policy aspect, and economic feasibility were checked. Firstly, technological feasibility was checked with three viewpoints such as appropriateness of technology development plan, possibility of technological success, and overlap with other similar programs. Secondly, policy aspect feasibility was checked with three viewpoints such as consistency with national policy, program management system, and risk factors. Thirdly, economic feasibility was checked. Previous and future funding details were analyzed to check proper funding, and to estimate average funding size and period. Budget details of similar programs were also analyzed to suggest proper funding size and period.

Finally putting all those together, future direction for the program was suggested.

Also the strengths and weaknesses of this evaluation methodology were described. The output or outcome of the R&D program is uncertain. So it is very difficult to infer the economic benefit of the R&D programs, especially for the basic research. Therefore it will be very difficult to promote national R&D programs in the area of basic research if the economic benefit is the key point of evaluation. By the way, this methodology will contribute to improve the planning of the R&D programs, and therefore to increase the possibility of success of the programs.

Thursday, 4 October 2018  
11:30–13:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 083 Promoting Resilience In Evaluation

## O 243 - Evaluating Resilience as a Component of Transition (Development) Impact

*B. Kolodkin<sup>1</sup>*

<sup>1</sup> European Bank for Reconstruction & Development, Evaluation Department, London, United Kingdom

**Background:** In 2016, EBRD modernised its longstanding concept of transition impact, the Bank's unique developmental focus. Previously, transition impact was focussed on transitioning centrally planned economies to market economies. Twenty-five years after EBRD's creation, the transition impact concept was believed to be a bit too limiting and perhaps somewhat out of date after EU expansion, the extended global financial crisis and other major disruptions to the status quo. The result was the creation of a new transition concept based on six qualities of a modern market economy. Those qualities are characteristics – competitive, well-governed, inclusive, green, integrated and ... resilient. In the new transition concept, resilience may be a primary or secondary transition (development) objective in EBRD projects.

**Rationale:** Evaluators must be prepared to evaluate resilience as both a project component and a larger theme. Evaluators at EBRD are in the process of developing the methodology, indicators, and products to provide stakeholders learning regarding contributions to resilience. This paper will present the issues the evaluators face in terms of methodological changes, evaluating projects today which were developed under the previous transition concept, data availability, organisational challenges, and understanding how private sector investments contribute to greater market resilience.

**Narrative/Objectives:** The paper will largely be an ex-ante assessment of undertaking resilience evaluation. EBRD's efforts in the area of resilience are largely based on financial stability, food security and energy security (with other potential secondary factors). A new compendium of indicators has been created by the Country Strategy and Results Management department to provide results in the area of resilience. Areas to be discussed in the proposed paper may include:

- Survey of other relevant techniques in evaluating resilience
- How to incorporate the new indicators into a Theory of Change
- How do these outcome-based projects contribute to larger impact (resilience) at the market, national or regional level
- Gaining buy-in from evaluators, project implementers, management and other stakeholders on this significant change

If data are sufficient and relevant, a mini case-study for Greece will be included. Greece recently became a temporary EBRD country of operations for 5 years because of the devastating Greek economic crisis and the effects on financial stability. One of EBRD's objectives is to build sufficient resilience to ensure appropriate withdrawal of Bank operations in a relatively short amount of time without causing instability.

**Value/Justification:** This proposed paper aims to:

- Provide insight into how international financial institutions view and contribute to resilience
- Contribute to the understanding of how lessons from evaluating resilience will aid building resilience in the European context
- Demonstrate how evaluators need to adopt new methodologies and techniques quickly in order to adapt to organisational changes
- Provide a relevant example, Greece, germane to EES and the local context

Thursday, 4 October 2018  
11:30–13:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 083 Promoting Resilience In Evaluation

## O 244 - Evaluation Advisers Able to Embrace Resilience – Insights from Department of International Development (UK) Experience

*B. Dillon*<sup>1</sup>

<sup>1</sup> DFID, Research and Evidence, London, United Kingdom

This Paper considers the skills and competencies necessary for evaluation professionals to remain relevant, resilient and cutting edge, in a variety of difficult environments and in face of new or changing demands. It focuses largely on the resilience of evaluation professionals themselves – their skills and competencies. However it does also consider resilience as an object of evaluation itself.

The Paper briefly looks at the range of contexts in which DFID works, and the demands this makes for a wider set of skills and competencies, including resilience, to be deployed. It then considers the evolution of the evaluation professional cadre in DFID since its inception in 2011, with particular reference to skills, competencies and challenges sought then and now, to show how priorities in professional skills need to change in response to demand. The Paper turns to consider the value added which evaluation professionals provide in building resilient societies drawing on their own skills, expertise and experience, and then at their contribution to the understanding of resilience as an object of evaluation.

The rationale for this Paper is based on the importance of the evaluation discipline building the necessary skills and competencies to remain relevant to face new challenges in an era where resilience is increasingly taking centre stage. Skills to ensure resilience are further hastened by new communication technologies and enhanced levels of capability being accessible by individuals. The objective is to share insights from the experience of a bi-lateral donor in the international development field with the wider evaluation community engaged in professionalization, with a view to influencing the level of resilience on the profession. This topic is highly relevant as professional skilling is an ongoing issue of discussion within this broad community. The input draws on DFID's own data and also the experience of a wider set of multi-lateral international development agencies. The input directly contributes to attempts to improve capability in the evaluation professional community. Resilience both in terms of skills and competencies, and in terms of measurement on the ground are addressed. This includes a gendered dimension, particularly relating to the contexts in which evaluation professionals have to operate. The topic is of public interest given the pressure and scrutiny on public services to operate value for money.

Thursday, 4 October 2018  
11:30–13:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 083 Promoting Resilience In Evaluation

## O 245 - Building a Culture of Transparency: Managing Evaluation to Support Resilience

*S. Farina*<sup>1</sup>

<sup>1</sup> Broadleaf Consulting, Evaluation, Vancouver, Canada

**Rationale:** Evaluation can contribute to alleviating a sense of crisis by providing framing and context for complex social issues, and important data essential to making good decisions. However, evaluations are often confidential for a variety of reasons. This presentation will address how evaluation can move away from producing unread or underutilized reports and into meaningful engagement, sharing information, and contributing to transparency in communities and institutions.

#### **Objectives sought:**

- Managing evaluation to support transparency and resilience
- Building a culture of transparency that supports civic engagement and good judgement

**Brief narrative and justification:** Innovative approaches to evaluation management offer opportunities to increase transparency, many of which are outside of traditional evaluation deliverables. Evaluations can be designed to meaningfully engage stakeholders, share information appropriately, support community ownership of and access to data, and communicate results beyond traditional reports. Resilience within societies and institutions depends on many factors and managing evaluation to support transparency during times of crisis is one way to support well-founded collective responses.

Critical thinking and good judgement are important for civic participation, and in political and economic systems, but applying these depends on a foundation of information. Increasingly, new trends in social media and political discourse are serving to limit information, present information out of context, or to present highly politicized information. The exercise of fair and transparent communication in evaluation is fundamental to supporting respectful, critical conversations.

Evaluation has the opportunity to transcend competing interests, and play a role in providing information in a way that is fair and transparent, so that civic and political dialogue can be based on a realistic understanding of the context in which interventions are taking place, the stakeholders, and the associated process and outcomes. This presentation will look at tools and strategies to manage evaluation in ways that are fair and transparent. The presentation will also discuss how a culture of transparency can be built into the process and management of evaluation.

Thursday, 4 October 2018  
11:30–13:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 084 Evaluation of Radicalization and Correctional Programs**

**O 246 - Evaluations in Prisons – Reflections on Evaluation Practice in a Challenging Context**

*W. Stuppert*<sup>1</sup>

<sup>1</sup> Camino gGmbH, Research associate, Berlin, Germany

Prisons and other institutions of the correctional system are a challenging setting for evaluations. Highly regulated access, strict internal procedures, and the overall coercive environment represent serious challenges to quality standards of evaluations such as the independence of the evaluation process and the protection of data subjects. At the same time, driven inter alia by a growing interest in radicalization prevention and deradicalization, prisons are of increasing importance to the European evaluation community.

How does the prison environment affect the quality of evaluations? What constraints are faced when different data collection methods are employed? What measures can be taken to mitigate the negative impact of this evaluation setting and maximize the specific opportunities that the prison environment offers to evaluators? What characterizes successful evaluations under these circumstances?

To answer these questions the contribution will bring together insights gained during the implementation of several evaluations of deradicalization programs and other interventions in German prisons with reflections on the challenges of doing research in correctional systems which can be found in the pertinent international literature.

These experiences and reflections will be discussed in reference to the quality standards of evaluation and ethical standards for social research. Problem areas that will be analyzed are inter alia institutional access, the organization of data collection in prisons and with the prison population, and the relationship between the researcher and imprisoned individuals. It is hoped that the presentation and discussion of the challenges faced and lessons learned while doing evaluations in prisons will contribute to the improvement of evaluation practices in challenging settings.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 138 The Transformative Power of Evaluation****O 248 - Challenges of Evaluating Impact of Sustainable Food Security and Undernutrition Intervention: Case Study of Bukoba Rural, Republic of Tanzania**N. Singh<sup>1</sup><sup>1</sup> University of the Incarnate Word, Nutrition, San Antonio, USA

**Rationale:** In Tanzania, undernutrition among children and mothers, especially in rural areas, is a common phenomenon. Evidence-based strategies for nutrition intervention mainly address *direct determinants* such as food/nutrient intake whereas, community based interventions tend to have higher efficacy when combined other *proximal determinants* such as women's economic and social empowerment. Women's education, access to economic means of production (such as land), ownership of assets and ability to purchase inputs (including from government support services) are very limited, which increases their income and food insecurity. It is relatively simple to measure the impact of direct determinants of undernutrition such as food/nutrition intake but becomes challenging measuring impact of proximal determinants. If you do not measure results, you cannot tell success from failure (World Bank, 2004) is true but presents challenge in small rural sustainable food security interventions as food is not stand alone variable.

**Objective:** the paper discusses the challenges of evaluating impact of both direct and proximal determinants in addressing undernutrition in rural Bukoba, Tanzania.

**Narrative:** In 2009, the women's perceptions of personal and social empowerment were assessed by a written survey, Women's Economic Development Personal and Social Impact Survey (WEDIS), administered in Swahili. Focus groups of 42 women representing 220 members of Bukoba Women's Empowerment Association (BUWEA), a registered Tanzania CBO were conducted to share the findings of the first Tanzania Mainland Nutrition Survey on children under five and mothers of child-bearing age. This survey documented a protein deficiency for this population. The focus group discussions led to identification of direct and proximal determinants of the problem of undernutrition. The identified *direct determinants* of undernutrition were lack of 1) availability; 2) accessibility; and 3) utilization of food whereas identified *proximal determinants* were women's lack of 1) sustainable purchasing power; 2) decision making power; 3) business acumen; 4) horticultural and technical assistance and 5) measuring impact of proximal determinants. These latter implied the lack of women's person and social empowerment in their family and community context. This data lead to a cooperative farming project creating a collective way to cultivate soybean to improve the quality of the existing diet (direct determinants) and an infrastructure to effect the women's economic sustainability and empowerment (proximal determinants). The following objectives directed the farming project: 1) To increase knowledge about the nutritional benefits of soybean consumption to various population groups (women, children and population with HIV-AIDS); 2) To increase soy consumption in the diets of the BUWEA members, thereby decreasing protein calorie malnutrition and improving overall health status; 3) To provide adequate technical assistance and training to the BUWEA members on the cultivation, expansion and commercialization (micro-enterprise) of the soybean in order to improve malnutrition, health and economic sustainability; 4) To increase the soybean crop yield for further expansion of the cultivation project for economic sustainability; and finally 5) To implement soy-processing machines/units such as vita-goat® to increase consumption of soy and create long-term sustainable economic activity (including micro lending) for the BUWEA members, their families and the broader community.

Thursday, 4 October 2018  
11:30–13:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 086 Turbulent Contexts and Evaluation 1

## O 250 - Evaluators for Planetary Health

A. Brouselle<sup>1</sup>, C. Butzbach<sup>2</sup>

<sup>1</sup> University of Victoria, School of Public Administration, Victoria, Canada

<sup>2</sup> University of Sherbrooke, Centre de recherche de l'hôpital Charles-LeMoine, Longueuil, Canada

**Rationale:** Entering into the XXI century came with the confrontation with a new context largely driven by environmental, social and economical challenges. New words also became part of our vocabulary: Anthropocene, mitigation, adaptation, resilience. The current state of planetary health inevitably leads to recognition of the failure of our current models of development and constitutes a call for each of us to profoundly revise our frameworks for action. As evaluators, while we make judgements on interventions, be they programs, policies, etc., we bear an important responsibility. We have the power to determine whether or not an intervention is good, whether it is arming or not. The first step is to accept that the way we conducted evaluations, as a group, until now, didn't lead us to the right direction. We now have to rise as a group and change fundamentally our practice to impact positively planetary health.

**Objectives:** In this presentation, we will (1) define the environmental challenges we face and how it is a multifactorial issue, (2) show that solutions are known and that we have to do is thinking how to implement them (3) discuss the change of role addressing the environmental threat involves for evaluators. Change is possible. But it requires that we make a commitment to new values and integrate them into our practice: (1) Sustainability: This will imply that in our evaluation we systematically assess in what manner the intervention will impact the environment, on the three elements, air, water and soil. We need to depart from silo thinking and embrace system thinking. We also need to systematically include long-term perspectives. (2) Equity: considering that environmental deterioration affects people all over the world, we should consider global impacts, and not just local ones. It also requires that we frame our results differently to make people engaged in this large agenda. This presentation is the result of a large scoping review of various fields of research on the environmental challenges we are facing: public health, sociology, psychology, among others.

**Justification:** This presentation aims at raising awareness but also at supporting evaluators willing to better address these challenges. The information we will provide will be pragmatic, giving guidance to all evaluators. This topic is of top priority internationally and is very well aligned with the priorities of action identified by the World Health Organization, The United Nations with the sustainable development goals, the Paris Agreement on climate change, among others. It will give direct recommendations to help evaluators develop new practices necessary for addressing the new environmental challenges.

Thursday, 4 October 2018  
11:30–13:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 086 Turbulent Contexts and Evaluation 1

## O 251 - Evaluation in Times of Crisis: Rethinking the Scope as well as the Methods in a Changing Environment

*G. Michailidis*<sup>1</sup>

<sup>1</sup> Aristotle University of Thessaloniki, Spatial Planning and Development, Thessaloniki, Greece

**Rationale:** Despite appearances, reassuring macro-economic signs and a general feeling that the crisis is over, challenges for the evaluation field do persist, either in the specific sectors, conjunctures and geographical areas, mostly hit till now, or as specific questions: "have we been useful in predicting the crisis? have our tools proved appropriate? shall we continue as if nothing has affected our scope and our methods? has our professionalism, standards and ethical norms been up to the needs and tasks?"

It is indeed possible that, especially due to our "anxiety" for evidence-based reporting and thus to the rush for quantification and for building elaborate indicator systems, we have, to a varying extent, overlooked the "strategic" aspect of our job. A rebalancing in the scope and methods of the evaluation seems opportune. The paper shall focus on the development policies, taking into consideration the "parallel" policies that frame or affect them.

**Objectives sought:** The paper soughts to establish:

- that all through the period of upgrowth of our job, we have managed to reach high standards of methods and specialized tools, capable to identify change and to provide pertinent suggestions as to how better elaborate policies and interventions
- that by doing so, we have proved ourselves useful in optimising policies in a generally stable economic, social and administrative environment
- that the "crisis factor", even if its most severe effects have been largely overcome, has in the meantime altered quite a lot of fundamental elements of the situation under evaluation
- that, consecutively, evaluating the starting point in the most correct way has become the most crucial part of our job
- that the evaluation of the policy tools is now obliged to take account of an environment of theirs quite unstable and sensitive to disturbances

**Brief narrative and justification:** The paper shall:

- assess the progress made in identifying challenges and opportunities as well as the still existing missing points in doing so
- try to define what is at stake in evaluating policies and interventions in an environment of crisis (incoming, "roaring" or coming off)
- try to induce possible appropriate changes or additions to our own tooling in order to understand and to confront them
- bring out new dilemmas and trends in professionalism, standards and ethical norms as well in what regards the role of partnerships and stakeholders.
- emphasize the need to interconnect the evaluation of a certain development policy with that of numerous "parallel" policies and, especially, to search for the possibility and the extent of synergies, both existing and searched for.

Thursday, 4 October 2018  
11:30–13:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 086 Turbulent Contexts and Evaluation 1

## O 252 - Evaluating Partnerships in Research and Innovation within a Mission-Oriented Policy Framework

*E. Amanatidou<sup>1</sup>, D. Cox<sup>1</sup>, C. Marzocchi<sup>1</sup>, D. Gagliardi<sup>1</sup>, M. Benaim<sup>1</sup>*

<sup>1</sup> University of Manchester, Alliance MBS - Institute of Innovation Research, Manchester, United Kingdom

Strengthening resilience in European societies is generally seen as the underlying goal of the current European policies in research and innovation. This has been made explicit in the Horizon2020 strategy promoting the concept of resilience whilst tackling demographic change, food security and the bioeconomy, energy efficiency, environmental and transportation issues, and security. These challenges have been referred to as 'wicked problems' for their complex nature and inter-relatedness which make solutions difficult to diagnose and implement (Boden et al. 2010). The presence of strong interdependences requires actions at the social, environmental and S&T level. These highlight the importance of multi-disciplinary research practices through multi-stakeholder approaches, multi-level governance and policy coordination across geographical boundaries and policy areas. Thus, they pose challenges in the design as well as the evaluation of policies and respective initiatives that address them (Amanatidou, et al. 2014).

One example of such policy initiatives are the transnational Public Partnerships (P2P) for research and innovation. These are partnerships among ministries or funding agencies across countries to deal research and innovation activities to deal with certain societal challenges. They include partnerships supported by the European Commission (such as ERA-NETs and Art 185s) as well as Member State-led initiatives (such as the Joint Programming Initiatives).

P2P are different from conventional competitive research programmes pursuing similar objectives: they present additional remits alongside their challenge-driven orientation. They are characterised by fluidity, with different types of membership and engagement. Their success often depends on the degree of trust between members as well as their long-term commitment. As transnational structures for collaboration they need to accommodate several national contexts (Lepori, et. al. 2014) regardless of the challenges of coordination and alignment of national research systems in terms of organisation, structures and funding (Nedeva, 2013). Consequently, the existing practices in evaluation and impact assessment based on research programmes evaluation fall short in addressing their special features.

The objective of the paper is to articulate a rationale for research evaluation under the 'challenge-driven policy orientation' (Foray et. al. 2012) and propose a specific framework for evaluation and impact assessment of P2Ps.

In the case of P2Ps, evaluation issues beyond efficiency and effectiveness become important such as network health and connectivity, while the issue of additionality spans various levels (national, trans-national, European). Impact goes beyond the scientific, technological and socio-economic domains. Particularly, policy/conceptual or structural impacts (Meaghar et al. 2008) become relevant changing the structures of national research and innovation systems. In this context, behavioural impacts have a different meaning.

Our approach builds on primary and secondary data elaborated under the ERA-LEARN 2020 project ([www.era-learn.eu](http://www.era-learn.eu)). This include statistical elaboration of data on FP7 ERA-NET actions (from eCORDA and an on-line survey with around 250 responses), 26 interviews of JPI members and two large-scale surveys of ERA-NET participants (600 responses).

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 087 Impact Evaluation In Complex Settings

## O 253 - Impact Evaluation of ACCRA Ethiopia Programme as a Moment of Triple Loop Learning – a Critical Contribution in Turbulent Times?

*J. Colvin<sup>1</sup>, M. Mukute<sup>2</sup>, M. Arranz<sup>3</sup>, H. Jeans<sup>4</sup>, M. Barihaihi<sup>5</sup>*

<sup>1</sup> Emerald Network Ltd, Learning Systems, Stroud, United Kingdom

<sup>2</sup> Social Learning & Innovation Ltd, Learning & Evaluation, Harare, Zimbabwe

<sup>3</sup> Oxfam Great Britain, Programme Quality Team, Oxford, United Kingdom

<sup>4</sup> Oxfam Great Britain, Resilience Knowledge Hub, Oxford, United Kingdom

<sup>5</sup> NDC Partnership, Country Engagement, Kampala, Uganda

The past decade has thrown up a wide range of innovative experiments in developing more resilient communities, organisations and governance systems, of particular relevance as we move into increasingly turbulent times. Evaluations can provide valuable opportunities to learn from these experiments, enabling both single, double and triple loop adjustments, adaptations and transformations. Triple loop learning (learning how to learn) is at a premium in turbulent times, in itself enabling the deepening of adaptive capacity and resilience. This paper draws on a recent impact evaluation of the Africa Climate Change and Resilience Alliance (ACCRA) programme in Ethiopia to draw out lessons about the role of evaluation in conditions of unpredictability and complexity, and to consider implications for the design and management of evaluations in turbulent times.

The paper considers ACCRA's learning in three interlinked areas: programme Theory of Change; evaluation methodology, design and management; and the broader role of the evaluation. As a theory-based evaluation, reconstructing ACCRA Ethiopia's Theory of Change in order to test its effectiveness and contribution to the intended transformation of governance systems lay at the heart of the evaluation design. The learning for ACCRA about theory of change lay in the way in which the ACCRA Theory of Change design was woven into the story of how transformational relationships developed and how these relationships contributed to slower and faster moving processes of innovation in governance practices. The evaluation review and its supporting learning process not only provided conceptual clarity and a way to talk within ACCRA about a complex set of processes and how these are interlinked, but also shaped a moment of triple loop learning with the potential to engender a step change in future transformational design. This clearly points to a broader lesson about key linkages between Theory of Change, appropriately designed evaluation and triple loop learning for deepening resilience.

To achieve this, the innovative evaluation methodology drew on process tracing, interweaving this with systemic approaches to evaluation including learning history. Furthermore, this innovation relied not only on methodological bricolage, but also on how this was embedded within a collaborative process of design and management involving the commissioning organisation (Oxfam GB), the broader ACCRA alliance of INGOs and the consultant team. At the heart of this process was a collaborative journey in which each party learned to let go of some initial expectations and preconceptions and came to understand the need to be flexible and trust each other to take joint decisions along the way. The broader lessons for the evaluation community challenge us to attend to personal resilience and to testing new assumptions and practices about inter-organisational working.

Drawing this learning together suggests new roles for evaluation in turbulent times, linked to moments of triple loop learning. Not only can these illuminate historical patterns of learning and adaptation pathways emerging across governance systems and other social configurations, but they can also provide the foundations for anticipatory learning about more transformational theories of change, in turn engendering more resilient pathways of development into the future.

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 087 Impact Evaluation In Complex Settings**

**O 254 - Building Livelihood Resilience in Rural Malawi: Constraints to Credible Impact Evaluation and Synthesis**

*J. Copestake<sup>1</sup>*

<sup>1</sup> *University of Bath, Professor of International Development, Bath, United Kingdom*

Poor, landlocked, predominantly rural and susceptible to climate change: how to promote greater economic security for citizens economically dependent on agriculture and natural resources dominates Malawi's development policy agenda. An additional feature of Malawi is a high level of dependence on international aid and its relative fragmentation: the site for endless initiatives aiming to promote diversification and commercialisation, climate smart agriculture, disaster preparedness and resilience. Strategies for sustaining smallholder livelihoods range from fertiliser subsidies and crop insurance to conservation agriculture and cash transfers. Yet for many the prospect of transcending precarious semi-subsistence farming remains remote. And while local and international development agencies self-identify as promoters of more positive outcomes, failures of coordination indicate that poor governance and aid management is also part of the problem. Credible, timely and generalisable feedback on the impact of different interventions is a necessary if far from sufficient condition for improving development practice to build rural livelihood resilience in Malawi. The paper draws on a range of case studies to review the challenges to providing such evidence and explores how far they can be addressed. Section 1 elaborates on the policy context of Malawi's quest to promote rural resilience, and on the shifting language of intervention strategies that informs it. Section 2 explores the methodological challenge to impact evaluation with reference to qualitative approaches to attributing impact. Section 3 explores the limitations of such evidence, trade-offs arising from reliance on more quantitative approaches, and scope for improving on both by mixing methods. The paper is relevant to professional evaluators, to those concerned with evaluating resilience as a policy goal, and to debate over scope for improving evaluation systems to improve development management and policy.

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 087 Impact Evaluation In Complex Settings**

**O 255 - Taking Impacts to Mixed Methods Evaluations in Ex-Post and Ex-Ante Settings**

A. Rowe<sup>1</sup>

<sup>1</sup> ARCeconomics, British Columbia, Maple Bay, Canada

Rapid Impact Evaluation (RIE) is a new evaluation approach for use in settings where it is otherwise challenging to assess impacts. RIE can be used to forecast expected impact, as well as to evaluate impact after implementation. It is nimble and low cost. Using RIE as part of mixed methods evaluations enhances the ability of evaluators to quickly and at low cost assess the direction and magnitude of impacts, including of complex and multi-system initiatives. Pilots in Canada, the US and SE Asia have shown RIE to be fit for purpose and provide valuable insights about how to tailor the method to different settings.

Built on a use-seeking framework RIE introduces three new methods: the scenario-based counterfactual; a simplified metric for scalar measurement of impacts; and an interest-based approach for using program stakeholders as experts. RIE triangulates judgments of three distinct groups of experts bringing knowledge of the intervention and of the science involved to the assessment of impacts. RIE is perhaps unique in being a practitioner developed approach that naturally infuses use-seeking approaches into its DNA.

This paper provides an overview of Rapid Impact Evaluation methods and provides a critical assessment of their use in national and international evaluations.

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 087 Impact Evaluation In Complex Settings

## O 256 - Counter-factual Impact Evaluation of EUREKA Network Projects and Cluster Projects – Strengthening R&I Policy Resilience through a Multi-Methodological Approach

*J.P. Kramer*<sup>1</sup>

<sup>1</sup> Prognos AG, Economy, Brussels, Belgium

Public budgets in many EU member states are under pressure and in particular R&I budgets had been decreased substantially in course of the financial crisis in many countries in Southern Europe, including Greece. At the same time, R&I support programmens, in particular collaborative R&D schemes with international partners, promise great additionalities and added value to the economy and society, thus contributing to more resilient economies through innovation.

Against this background, the **objectives** of the **Counterfactual Impact Evaluation** of EUREKA Network and Cluster Projects was to (a) assess the investment volumens mobilised through the support, (b) understand the main motives for participation in EUREKA (compared to other EU R&D schemes), (c) analyse the impact of the support on turnover and employment, and (d) explore the main success factors for the projects and the exploitation of results from public R&D schemes.

The **CIE of EUREKA Network Projects and Clusters Projects** used an in-depth evaluation using a multi-methodological approach. The aim was to both provide quantitative evidence about the impact of the international R&D collaboration schemes but to open the black-box about the context, causal mechanisms and longer-term strategic value of the support. The assessment was carried out by a consortium led by Prognos AG and its research partners, Joanneum Research and Professor Cincera (Université Libre de Bruxelles). All participants in projects which received an EUREKA label in the period of 2001–2015, either as a Network or as a Cluster Project, were studied. The empirical work was based on a European wide survey, an econometric analysis, and in-depth case studies.

The CIE approach utilised was **novel** in its integrative nature of approaches, paying highest attention to triangulation and the combination of quantitative and qualitative evidence. This was achieved by both designing the evaluation around this ambition and by combining different data sources, including explorative interviews, project-level data (monitoring data), a survey addressing both supported and non-supported firms under EUREKA, a counter-factual impact evaluation using a DiD-approach and finally comparative case studies.

The analysis confirmed, amongst others, that **annual turnover of participating firms develops significantly better** than turnover of the non-supported firms over time. R&D in EUREKA projects is meant to be “near-to-the-market”. Therefore, firms should be able to exploit the R&D results quickly. As the econometric assessment has shown, the average annual turnover growth is in general positive and significant for participants of both Network and Clusters Projects in comparison to the firms not receiving support (control group). For example, one year after the end of projects, Network and Cluster Project participants showed an additional annual turnover growth of 15% and 13% compared to non-funded firms. Besides an overall positive impact of project participation on turnover and employment, which is significant compared to the non-funded firms, further effects on the economic performance of EUREKA participants were collected by the survey. Besides an increase in turnover (73%), EUREKA projects are expected to support the entering of new (local or foreign) markets (69%), improve (local or foreign) market shares (68%), or increase exports (67%).

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 088 Innovative Approaches and Methods for Unpacking Complex Causality**

**O 257 - A Perfect Match: A Demonstration of Pattern-Matching Techniques to Test Complex Program Theories**

E. Raimondo<sup>1</sup>, J. Vaessen<sup>2</sup>

<sup>1</sup> World Bank Group, Independent Evaluation Group, Silver Spring, USA

<sup>2</sup> World Bank Group, Independent Evaluation Group, Washington, USA

Pattern-Matching is a well-established method of causal inquiry in evaluation. Originally developed by Trochim in the 1989, it recently resurfaced as a useful theory-testing approach that lends itself to case-based and configurational causal analysis. At the interstice of theory-based and case-based evaluation pattern-matching has the potential to improve the rigor and transparency of evaluation findings when studying complex processes of change. This demonstration session will take the learners through a real-world example of the application of pattern-matching to the evaluation of the World Bank's catalytic role in developing carbon finance. The 10 steps of pattern-matching will be covered in a dynamic manner. Caveats and challenges in applicability will also be discussed. By the end of the session, the participants will be ready to become to achieve the "perfect match" between theory and empirics.

**Theoretical and Methodological Relevance** Since the ground-breaking report on broadening the range of designs and methods for impact evaluations (Stern et al. 2012), the evaluation community has been experimenting with new qualitative and case-based methods of causal inference that are grounded in a detailed understanding of theories of change and rival explanations. Chiefs among these newly introduced approach are Process-Tracing and Qualitative Comparative Analysis. The specific example of Pattern-Matching that will be discussed during the session is at the juncture between the two. Sixteen cases of process-tracing (without Bayesian updating) were conducted and compared in a configurational manner using necessity and sufficiency analysis. This approach presents a number of advantages in strengthening the quality and rigor of theory-based evaluation approaches: (1) it improves the specificity and clarity of theories of change and formalizes the "theory testing" part of the evaluation, which is often underspecified; (2) it provides evaluators with guidance about how to collect data consistently across cases and how to assess the probative value of the evidence collected; (2) it reduces confirmation bias; (3) it enables mid-range generalizability; and (4) rigorously prepares the data for formal testing via QCA. Yet, the true theoretical, methodological and practical value of the approach can only be revealed after its passes muster with the community of "real-world" evaluators. The examples of application of pattern-matching in the evaluation of the World Bank Group's support to Carbon Finance in its client countries contribute to establishing the relevance and feasibility of the approach.

**Value to the field of evaluation** While Process-Tracing and QCA are particularly *en vogue* in the evaluation community (there were no less than 15 presentations on this topic at the previous European Evaluation Society's conference), real-world evaluation applications remain rare to date. In the evaluation of the World Bank Group's support to carbon finance, the method was refined to adapt to real-world constraints. The true value of innovative methods becomes evident only when a critical number of applications have taken place. This demonstration session contributes to this collective endeavor of building a body of evidence around the application of these innovative methods of causal inference.

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 088 Innovative Approaches and Methods for Unpacking Complex Causality**

**O 258 - Applying the Principles of Outcome Harvesting for Monitoring and Evaluating Development Interventions**

*R. Wilson-Grau*<sup>1</sup>

<sup>1</sup> Ricardo Wilson-Grau Consultoria em Gestao Empresarial Ltda, Independent Consultant, Rio de Janeiro, Brazil

Are you a grant maker, manager or evaluator who is using or plans to use Outcome Harvesting for monitoring and learning (M&E) or for a developmental, formative or summative evaluation? Do you experience the need to customise the six Outcome Harvesting steps to meet your specific needs? If yes, then you are the audience for this session. I will exemplify how the ten underlying effectiveness principles can guide you in creatively using the approach.

Outcome Harvesting has proven most useful for a wide diversity of interventions operating in dynamic contexts with considerable uncertainty about what will be achieved until outcomes emerge. Furthermore, the participatory, qualitative Outcome Harvesting process itself is complex, with design decisions being made step by step. Expectedly, this has presented a major challenge to remain true to Outcome Harvesting. When do the changes made in applying Outcome Harvesting cross the line and the evaluation stops being Outcome Harvesting?

The lines are not easy to draw. Outcome Harvesting turns the normal evaluation process on its head in many ways. For example, a common expectation of evaluation commissioners is to contract an evaluation and then stand back and wait for independent, external experts to collect data and render their judgments. This contrasts with the participation in decision-making and information collection that are at the core of Outcome Harvesting. Consequently, right from the beginning of an outcome harvest, you are under pressure to adapt to circumstances. And, the adaptation and customisation continue through the harvest.

As Outcome Harvesting became mainstream, I began seeing signs of misunderstanding on the part of users and evaluators alike. I became concerned if not anguished to find almost as many misuses and abuses of the Outcome Harvesting approach as creative advances in developing it further. The pressures created by diversity, uncertainty, constant change, participation and unpredictability, coupled with the usual demands of heterogeneous primary users, were undermining fidelity to the concepts behind each Outcome Harvesting step. I concluded that one of the fundamental reasons, and perhaps the principal one, was that I had not identified and much less explained the underlying Outcome Harvesting principles.

Therefore, in 2015, I identified the core principles that I and co-evaluators have used to maintain the fidelity of the approach as we customise the six steps to specific needs and different contexts. I began in 2016 a process of making these implicit principles explicit, first for myself and co-evaluators and then for others through the conferences of the American (AEA), African (AfrEA), ReLAC, IDEAS, National Evaluation Capacities and the EES conferences. That same year, I contributed a chapter "Outcome Harvesting Evaluation – Practical Application of Essential Principles" for Michael Quinn Patton's *Principles-Focused Evaluation: A GUIDE*, published last year. Now, I have written a book myself, *Outcome Harvesting for Monitoring and Evaluation: – Practical Applications of Essential Principles* that will be published by IAP to coincide with the EES 2018 Conference. This session intends to contribute to another solid step forward.

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 088 Innovative Approaches and Methods for Unpacking Complex Causality**

**O 259 - Applying Outcome Evidencing to Unpack Complexity and Contextualise Results in a Livestock Project in Zimbabwe**

*H.J. Garcia<sup>1</sup>, R. Azuba<sup>2</sup>, D. Chikodzore<sup>3</sup>*

<sup>1</sup> Food and Agriculture Organization of the United Nations, Office of Evaluation, Rome, Italy

<sup>2</sup> Independent Consultant, FAO Office of Evaluation, Kampala, Uganda

<sup>3</sup> Independent Consultant, FAO Office of Evaluation, Mutare, Zimbabwe

The paper will discuss the use and lessons learned in applying Outcome Evidencing methodology (Paz-Ybarnegaray and Douthwaite 2016). It will elaborate on the experience of using this method in evaluating a livestock project in Zimbabwe's Matambeland Province.

The project's objective was to encourage smallholder livestock farming as business through improving animal health, market linkages and improving policy for livestock. The interconnectedness of the project's intended results would contribute to improve resilience of communities, in the two dry land districts of Lupane and Nkayi, which is commonly prone to drought and surges of livestock diseases.

Outcome Evidencing was used in the initial stage of the evaluation to unpack the complexity of the intervention and its implementation approaches and contextualize the results achieved. A workshop was held with 25 "change agents", who were representatives from the government and two implementing partner organizations, coming from the national, provincial and district levels. The participants identified six "Areas of Change" that were used to extract 150 results (output and outcome level) stemming from the projects. Through several participatory analyses, the results were condensed into 78 outcomes level results. These outcomes were used to build six multi-causal diagrams. After, a "critical chain of results" was selected in each diagram. Each of the 78 outcomes and linkages in the multi-causal diagram served as testable "evidences". Some of these were tested and substantiated through field visits, focus group discussions and interviews. As a modification of the methodology, a Strength-Weakness-Opportunities-Threats (SWOT) analysis was conducted for each "critical chain of results". This added an additional source of verifiable evidence for the outcomes.

Overall, Outcome Evidencing provided a systematic entry point to the project evaluation and increased the engagement of the stakeholders. The methodology helped aggregate, validate and construct causal linkages around seemingly dispersed and independent project components, in tandem with other evaluation methods. The outcomes that were defined in the workshop were used for further triangulation and for snowball method. The multi-causal diagrams and Strength-Weakness-Opportunities-Threats analysis of "critical chain of results" was used to direct some of the field validation exercise through field visits, providing for a more efficient use of evaluation resources. Since the project's components were being implemented by "change agents" at various government and implementing partner levels (i.e. national, provincial and districts), the workshop contributed to additional cross sharing and added to the utilization of the evaluation process.

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 089 Evaluating Educational Programs**

**O 260 - Play It Different This Time, Sam! Involving Teachers and Educators in Evaluation of Drop-Out Prevention Programs in Schools**

*L. Tagle<sup>1</sup>, S. Celano<sup>2</sup>, V. Fini<sup>3</sup>, S. Pirozzi<sup>4</sup>*

<sup>1</sup> *Italy's Presidency of the Council of Ministers, Evaluation Unit - Department for Cohesion Policy, Roma, Italy*

<sup>2</sup> *Independent Evaluator, Independent Evaluator, Avellino, Italy*

<sup>3</sup> *Independent Evaluator, Independent Evaluator, Massa, Italy*

<sup>4</sup> *Independent education consultant, Independent consultant, Naples, Italy*

The paper discusses how evaluation may facilitate social change by promoting front-line actors' access to and process use of evaluation – above and beyond transferring evaluation results to program funders and coordinating agencies. We base our claims on our experience evaluating "Play it again, Sam!" (Sam), a program for prevention of school drop-out and of educational poverty which has operated for more than twenty years in Turin, a middle-sized city in Northern Italy.

Called in by Fondazione per la Scuola (FpS, the private foundation which provides most of Sam's funds), we evaluated the three-year pilot project which, starting in 2015, has deeply innovated Sam. Sam used to focus on individual at-risk students in their first year of junior high school. Now, this focus has shifted in three ways. First, the target of the "new" Sam is the teachers' team responsible for each selected class. Second, eligible classes are now both junior high and primary schools. Third, Sam activities now support the whole class as a group (rather than at-risk individuals). Other changes pertain to the length of support (extended from one to up to three years), the requirement that teachers and educators together plan Sam's activities for three years (rather than each year), and that educators participate in teachers' meetings, including the meetings aimed at assessing students' learning achievements. These changes aim at contaminating junior high teaching practices with primary school teachers' and educators' knowledge, competences, practices, and values. Our evaluation team was retained to help establish the monitoring system and to support the program community via on-going evaluation. Our mandate required closely scrutinizing the innovations introduced in the program. We interpreted this mandate by closely involving the main adult actors in monitoring and evaluation activities. We engaged representatives of teachers, educators, program funders, and the city government in developing the metrics and indicators for the monitoring system. This led to shared identification of outcomes which were not originally envisaged by planners.

We discussed most evaluation tools, such as surveys and focus groups, with educators and teachers. In some cases, we implemented these tools with the help of respondents. Involving front-line actors appeared more effective in transferring and debating preliminary results than periodical reports and meetings with program funders and stakeholder representatives.

Apart from increasing the validity of our findings, involving front-line workers promoted acceptance and use of evaluation results in a situation where adults (funders, the city, the teachers, school managers, and educators) need to alter their practices to face increased difficulties in both poor and middle-class neighborhoods while navigating the straits of reforms and ever-shrinking resources. For example, educators used involvement in the evaluation to achieve increased legitimization vis-à-vis program governance organisms and funding entities. They had already created their own coordination group, which interacted with the evaluation team. The final report by the coordination group included concepts and information they derived from their experience participating in the evaluation.

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 089 Evaluating Educational Programs

## O 261 - Sustainable Development in The Context of Vocational Competencies and the Goals of Agenda 2030

*M. Rökköläinen*<sup>1</sup>

<sup>1</sup> *The Ministry for Foreign Affairs of Finland, Development Evaluation Unit, Helsinki, Finland*

Sustainable development is a key competence for lifelong learning as well as a central part of vocational competence. The Finnish Education Evaluation Centre conducted an evaluation of the sustainable development learning outcomes of vocational upper secondary qualifications in spring 2017. The evaluation was addressed to all providers and participants consisted of a total of 5000 students studying for different qualifications.

The aim was to evaluate how well the ecological, social, cultural and economic sustainable development objectives have been achieved. In addition, the aim was to evaluate how the operations and teaching of education providers relate to learning outcomes. The evaluation aimed to produce information that education providers could utilize in the development of their own activities and teaching as regards the sustainable development objectives. This evaluation is also linked the sustainable development goals of Agenda 2030.

The evaluation was conducted entirely electronically. One of the objectives in the evaluation was to develop the electronic assessment test so that it could be used to assess the key competences for lifelong learning. The evaluation consisted of a test for learning outcomes as well as a self-assessment for education providers and for teachers.

The evaluation revealed that monitoring learning outcomes and purposefully focusing on individual factors that affect learning outcomes makes it possible to ensure equal opportunities for all students to learn and achieve a level of competence corresponding to sustainable development objectives.

Learning outcomes are not affected by how advanced the education provider's strategy is, but practical operations in regard to sustainable development at the educational institution seem to be connected to the competence of students.

The importance of teaching in regard to learning outcomes is emphasized particularly when the student has not learned about sustainable development in connection with practices at home or hobbies.

Competences required by the working life can be improved by implementing the teaching of sustainable development in a more employment-oriented manner and by developing the content of instruction so that the ecological, social, cultural and economic dimensions of sustainable development are balanced.

The evaluation showed that an electronic assessment test can be used in the assessment of the key competences for lifelong learning by focusing on the assessment of knowledge-based and operational competence and supplementing it with a student self-assessment. In order to ensure the relevance of the assessment, practical tasks specific to each field of qualification should be further developed. Ensuring that the results can be utilized requires the development of new feedback practices suitable for electronic assessment and methods for ensuring students' participation in the self-assessment process.

Finland is very committed in realizing the Agenda 2030 and is planning cross-sectorial evaluation of the implementation. In my presentation I will also reflect the evaluation results of education sector in the point of view of Agenda 2030 and forthcoming follow-up and evaluation.

Source: Rökköläinen, M., Metsämuuronen, J., Holopainen, J., Hievanen, R.: Sustainable development in the context of competence, teaching and the operation of education providers in vocational upper secondary qualifications (FINEEC 2017)

Thursday, 4 October 2018  
15:00–16:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 090 Evidence for Social Justice

## O 262 - How Can We Use Evidence Better for Achieving Social Justice? Lessons from Oxfam's Evidence for Influencing Conference

*L. Guijt<sup>1</sup>, I. de Goede<sup>2</sup>*

<sup>1</sup> Oxfam Great Britain, Research and Publishing, Oxford, United Kingdom

<sup>2</sup> Oxfam Novib, Evaluation, The Hague, Netherlands

This paper shares recent lessons from Oxfam on how to use evidence when seeking social justice. Since Oxfam started 75 years ago, its influencing work has evolved in response to issues and context. Oxfam staff use trends and horizons scanning to inform strategic positioning, audience research to develop global campaigns and evaluative research to enhance the effectiveness of programmes. In addition, staff conduct many thematic policy studies and stakeholder analyses as input for advocacy strategies to change the terms of debate.

Over the years, Oxfam staff have compiled business scorecards, global indices and tax monitors. Others have researched statistics and generated 'killer facts'. Some have undertaken impact studies, while others have done deep dives to understand the whys of impact. And some have collected hundreds of stories of personal experience to detect options for policy and practice. Many of these efforts have been collaborative, with CSOs, other NGOs, universities and think tanks. In the process, Oxfam has invested considerably in being evidence-informed. It has gotten it right – when the data is robust, packaged well and hits home. And sometimes it clearly hasn't. So how can Oxfam do better?

In October 2017, the authors organised an evaluative conference at Oxfam to take stock of its use of evidence in influencing for social justice. Such change is a journey during which many kinds of evidence play a critical role. Oxfam needs to know *what* needs to change, whether it be government policies, company practices or public attitudes and behaviours, as well as innovative solutions. Oxfam also needs to know *who* has power to make the change and *who/what* influences them. If it's the private sector, which company and why? If it's the government, which department, parliamentarian or civil servant and why? If it is the public, which ones and why? And Oxfam needs to know *how best* to achieve change including which influencing strategies and tactics work to shift policies, practices and norms, where, when and why? All these questions require research and other forms of generating evidence.

None of the dozens of experiences shared in 2017 were a walk in the park. Everyone had a first-hand experience with the *tensions and challenges* of embedding a robust evidence base in the influencing journey. Most were painfully familiar with having too few resources such as money, time and people, to get to robust insights. Everyone juggled the needs of influencing tactics with research integrity. Researchers' desire to share every nuance struggled for space alongside those who need simple messages and simple asks. Global advocacy wanted to emphasise different findings than country-specific efforts.

This paper will describe the main insights around three themes: (1) assumptions made about how change happens; (2) the challenges and implications of populism for using evidence for influencing; and (3) the efficacy of different kinds of evidence to shift the terms of debate and policy decisions.

Thursday, 4 October 2018  
15:00–16:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 090 Evidence for Social Justice**

## **O 263 - Evaluating Unintended Effects and Adaptation in Countries Affected by Fragility, Conflict and Violence**

*S. Tamondong*<sup>1</sup>

<sup>1</sup> UN Women, GEAC, New York, USA

The Role of Women in Evaluating Conflict and Fragile States

**Rationale:** Most often women are neglected during evaluations in countries affected by fragility and conflict. Yet, they are mostly affected even unintentionally. Thus, it is important to evaluate their situation as women in fragile states are even more marginalized and vulnerable to poverty.

**Objective:** To raise awareness and understand women's roles and predicament during conflict and in fragile states; and

To learn what are the gender sensitive indicators in evaluating unintended effects and adaptation in countries affected by conflict.

Women are particularly marginalized and under represented in the evaluation of fragile states and in post conflict societies. Around half of the world's poor live in fragile states and majority of them are women. As fragile states make people poor, fragility make people more at risk of poverty, making it harder for them to escape poverty. Women living in fragile states are even more marginalized and vulnerable to poverty. Six out of ten of the world's poorest are women. (UNDP Gender and Poverty Reduction report). The risks of poverty are even greater for women living in fragile states. Fragile states reinforce gender inequality. Fragile states turn to a patriarchal world view, traditional, religious and customary laws due to state weaknesses and ineffective institutions. Fragile states lack democratic accountability, which is a challenge to women. One in 3 women worldwide are victims of domestic or sexual violence in their lifetime, and women in fragile states are particularly at risk of being victims of violence. There is a higher level of societal violence and acceptance of violence in fragile and conflict areas. (WHO Report 2013 Violence against Women..." Fact Sheet 239.

If the need to promote women representation is not structurally addressed, early on, in the rebuilding of nations, as a matter of priority, rather than as add on, women's contribution in political and social processes, will be severely undermined. It would be a lost opportunity.

Women champion change beneficial to everyone – they raise social and practical issues, related to health, children, education and livelihood, water use and sanitation, issues women encounter in daily life- which men in power may not see easily, or at all. Strategies to address and improve accountability and governance must build women's participation. Women can be agents of change promoting social justice. Ways to involve women in formal and informal negotiations in transitional democracy and political settlements must be found. Women must be allowed to define and experience justice. Psychosocial support to women and girls in the aftermath of conflict is essential for lasting peace and development. There is a need to unlock gender equality in fragile states and conflict affected situations. Thus, in evaluating unintended effects and adaptation in countries affected by fragility and conflict, the role of women is imperative.

ST/ March 15, 2018

Thursday, 4 October 2018  
15:00–16:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 090 Evidence for Social Justice**

**O 264 - Values, Judgment and Reasoning in an Era of Evidence**

*K. Hassall*<sup>1</sup>

<sup>1</sup> *The Australia and New Zealand School of Government, Evidence and Evaluation Hub, Sydney, Australia*

Evidence doesn't speak for itself. Meaningful evaluation evidence is the synthesis of empirical information with relevant values; a product of evaluative reasoning leading to warranted judgments. In this era of evidence, there are debates about what constitutes valid knowledge for informed decision-making.

I argue that the demand for useful evidence requires the evaluation profession to develop better concepts and practices for engaging with values, and more explicit practices for reasoning and evaluative judgment.

Linking theory and practice, this paper draws on recent work on judgement by philosophers, combined with work on values by social psychologists as well as the decades of work by evaluation theorists on the role of values and evaluative judgment. It proposes a way for evaluators to think about values, judgment and reasoning that can help us to improve our practices of generating evaluation evidence.

For decades, key evaluation thinkers have advocated that we practitioners pay greater attention to our practices of judgment and be more explicit about values. However judgment has been considered suspect by some disciplines, because it is perceived as not objective. Thus we have relatively fewer formal methods and guidance on the use of values and judgment, compared with the level of attention paid to research methods and evaluation use. We still don't have widely practised methods for evaluative reasoning and expressing warranted judgments – compared with other areas of evaluation methodology.

My aim for this paper is to help evaluators put the advocated theory into practice, by providing a conceptual framework to engage with values that allows us to be more explicit about evaluative judgement in our evaluation practice.

Emeritus Professor Saville Kushner has described evaluative enquiry as “a process for arriving at judgments about public value, in such a way that reveals the nature of the public”. By understanding new ways to think about, and work with, values and judgment, we will be able to make more meaningful evaluative judgments in this era of evidence.

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 091 Humanitarian Interventions and Evaluation

## O 265 - Engaging with the Affected Population in the Evaluation of Humanitarian Responses

*R. Alsalem*<sup>1</sup>

<sup>1</sup> *Independent, Consultant on forced displacement and humanitarian response, Brussels, Belgium*

Engaging the affected population when conducting an evaluation of a humanitarian response is not only one of the most important aspects of that evaluation but also one of the most difficult. Access to this population has also become more challenging given difficulties of access; insecure environment for the organisations and institutions involved in the response, including the evaluation. Finally, if approached incorrectly, conclusions that are made within the course of an evaluation can be misleading and incorrect.

In the following paper, I will be looking at some of the lessons learned from the evaluations I have conducted when it comes to involving the affected population more effectively in all of the phases of an evaluation:

- a) Preparation
- b) actual evaluation and
- c) follow-up to an evaluation.

To mention a few fundamental aspects:

- a) Preparation of the evaluation
  - Ensuring that you have a full understanding of the profile, size, and attitudes of the affected population; leadership and representation structure; political interests;
  - Understanding in detail how the affected population has been involved in past evaluations, and what their reactions were to the process;
  - Planning carefully the communication about the upcoming evaluation to reduce manipulation or misunderstandings by others;
  - If translation is needed, spend time hiring the right translators (with background checks to avoid closing someone that is politically affiliated or that has abuse of authority issues) as well as train them on the type of translation that you;
  - Refraining from leaving it up to community leaders to decide on who should be interviewed. Involve also civil society, international organisations and the authority. Make sure your samples are age, gender, and diversity sensitive. Do not leave out minorities.
- b) Conduct of the evaluation
  - Combining group discussions with one on one interviews to double-check information and to provide different confidential opportunities for persons to share information;
  - Spending generous amounts of time explaining the purpose of the interview, as well as the evaluation overall. Allow sufficient time for questions and answers by the interviewee to make sure it has been understood;
  - Manage confidentiality and privacy carefully. If you can not be fully alone with the person(s) that are being interviewed, refrain from posing sensitive questions that may put the interviewee at risk for answering them. Knowing where your "interviewees" gather or like to go (for example, health clinics'; schools, etc.) and spending time there will provide you with ample opportunities to have one to one conversations in a safer space;
  - Observation is as important a tool as is conversation. Observing the persons you are interacting with as well as your surroundings is key which will give you insights as to the next steps you should take.
- c) Follow-up
  - If in the course of the evaluation, individual cases or groups of persons with specific needs have come to your attention, share it with those following up in a safe manner;

- Ask to know from your client, how the results of the evaluation will be shared with the community;
- Recommend in your report a communication and follow-up plan.

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 091 Humanitarian Interventions and Evaluation

## O 266 - Use of the Evaluation for the Programme on Violence Against Children in School in Jordan

*H. Tsuruoka*<sup>1</sup>, *M. Homsj*<sup>2</sup>, *Y. Matsuda*<sup>3</sup>

<sup>1</sup> UNICEF Jordan, Planning- Monitoring & Evaluation, Amman, Jordan

<sup>2</sup> UNICEF Jordan, Child Protection Section, Amman, Jordan

<sup>3</sup> UNICEF Jordan, Planning- Monitoring and Evaluation Section, Amman, Jordan

UNICEF has commissioned an evaluation of the programme on violence against children in school in Jordan (Ma'an Programme). The programme has been in place since 2009 and went through the Syria Crisis. Jordan today hosts the third largest population of registered Syrian refugee after Turkey and Lebanon. The Syria Crisis has stressed the already burdened education system, such as overcrowded classrooms and high student-teacher and -counsellor ratio. With the historical background of hosting large Palestinian registered refugee population, building resilience and social cohesion among different segments of the society has been a persistent critical issue to address in Jordan. Given this context, this evaluation has contributed to generate evidence to inform the decisionmakers.

The evaluation used a mixed-methods approach to evaluate the programme against OECD DAC criteria. The quantitative data was retrieved from database of the Monthly On-line Survey System administered by the Ministry of Education. For primary data collection, a range of qualitative data collection methods such as key informant interviews, focus group discussions, and field observations were used. The use of Mixed-Methods for data collection enabled the evaluators to offset the weakness of using a single method and thus helped better validation of information, perceptions, and opinions.

The entire evaluation was undertaken in a participatory manner, which is demonstrated through active involvement of key stakeholders for their inputs at critical stages such as evaluation design, tools development and application. In addition, the beneficiaries, in particular children and parents, were also consulted for their views on programme strategies, implementation, achievements and recommendations for future of the programme.

The evaluation concluded that the Ma'An Programme was an undertaking of national importance to address the deep-rooted problem of violence against children in Jordan. The programme has created a national acceptance and a momentum against violence against children in schools, with a balanced, inclusive, gender sensitive and human rights based approach. The achievements realised at astoundingly low costs is commendable.

It was indicated that the approaches and strategies adopted for implementation were seriously challenged due to unforeseen external factors and internal oversight. The influx of Syrian refugees resulted in a shift of UNICEF's resources towards humanitarian assistance and hence the Ma'An thrust, efforts and resources were diluted within UNICEF Jordan and the Government of Jordan. The evaluation indicated that despite all these challenges UNICEF Jordan and Ministry of Education remained determined in the implementation of the programme.

Based on the evaluation recommendations, UNICEF Jordan together with line ministries and other stakeholders has launched a multi-year strategic plan to eradicate violence against children in all settings, including schools.

Unpredictability of Syria Crisis and its impact to affected countries gives a unique role to the programme evaluation. In the humanitarian setting, development programme is often paused with emerging priorities and adversely impacted by the emergency. Yet, programme evaluation provides an opportunity to demonstrate importance of resilience building by informing the effectiveness and long-term impact of the programme. Evaluation can objectively identify gaps in programme and at the same time help to clarify priorities in unpredictable circumstances.

Thursday, 4 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 092 Environmental Evaluation**

## **O 267 - Evaluation of Environmental Protection Processes. Application in a Major Industrial Installation in Greece**

*I. Panagopoulos<sup>1</sup>, A. Karayannis<sup>1</sup>, G. Gouvalias<sup>2</sup>, H. Karoukis<sup>1</sup>*

<sup>1</sup> SYBILLA Consulting Engineers Ltd, Environmental / Safety, Athens, Greece

<sup>2</sup> SYBILLA Consulting Engineers Ltd., Safety, Athens, Greece

**Rationale:** This paper presents carried out work on the realization path, mode selection. some of the results and the self-evaluation of Environmental Protection Processes as applied in a Major Industrial Site, namely the self-evaluation of Preliminary Orientation Environmental Liability Assessment Study (POELS) of a full industrial scale metal treatment installation.

**Objectives sought:** The ultimate goal of this work is to illustrate the role of the self-evaluation of Environmental Protection Processes in Environmental Liability issues in order to identify/assess/evaluate major risks at the facility and risk mitigation measures.

**Brief narrative and justification:** The steps that were undertaken during the Environmental Protection Processes at hands (POELS) consist of Preliminary Risk Identification, Preliminary Measurements and Preliminary Risk Assessment, Preliminary Identification and Assessment of Risk Mitigation, and Self Evaluation of POELS.

A number of significant processes/activities that may result in a risk to environmental receptors are analyzed. Risk classification criteria, were designed to reflect the critical levels of risk appropriate to the installation from these activities. Risk ratings were applied to each risk for severity and occurrence and a risk score was calculated. In that context a self-assessment/audit of the project with regard to its planning, implementation, results is carried out, It involves the objective assessment of the ongoing project, its design, implementation and results. It aims to improve policy and practice, Enhance accountability, determine relevance and fulfilment of objectives, as well as efficiency, effectiveness, impact and sustainability of a project. Self-Evaluation/audit also makes statements about the relevance of planned outputs/outcome, the achievement of the outputs/outcome (effectiveness), the efficiency of the project, the sustainability, the impact and makes recommendations on the further development of the projects evaluation

Environmental Protection Processes Project Evaluation Criteria follows the OECD-DAC: A-Relevance/Appropriateness: B-Connectedness (and coordination): C-Coherence: D-Coverage: E-Efficiency: Were the results delivered in the least costly manner possible?, F-Effectiveness: G-Impact:.

Overall, the evaluation review found that POELS offered a good quality induction programme for first-time study related to Environmental Liability legislative needs. There were some aspects of the programme which required fine-tuning to maximize their effectiveness, but no major changes were recommended by the review. The review found that POELS provided a good platform for inducting first-time principals across relevant preliminary settings drawing a baseline framework on the use of detailed strategies to investigate causal pollutants sources attribution related to Environmental Liability legislative needs.

Thursday, 4 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 092 Environmental Evaluation**

**O 268 - Becoming Bigger, Better, Smarter: A Review  
of the Evaluability of Climate Change Proposals Presented  
to the GCF**

*J. Puri*<sup>1</sup>

<sup>1</sup> *Green Climate Fund, Independent Evaluation Unit, Songdo, Republic of Korea*

The presentation will focus on key learning from an analyses of GCF climate change proposals. The GCF aims to be transformational and aims to move economies to low-carbon, high resilience pathways. To what extent to proposals and projects that are being supported by the GCF incorporate this and inform this overall objective in an evidence based way? With two other co-authors, the presenter will summarize some key findings and lay out some learnings for programs and projects. The paper will analyse the strength of evidence, theories of change, causal pathways, impact potential and innovativeness of these projects and then present a stoplight of how well the portfolio is performing.

Thursday, 4 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

S 092 Environmental Evaluation

## O 269 - Re-Thinking Hierarchical Evaluation Systems Towards Evaluation Theory of Change that Defines Sustainability in African: Using the Psychological Perspective to Demonstrate

K. Odhiambo<sup>1</sup>

<sup>1</sup> University of Nairobi, Psychology, Nairobi, Kenya

**Rationale:** Designing national programmes in evaluative terms of “power” and “influence” of who is ‘doing well and better,’ ‘who has made big gains,’ ‘who is progressing faster,’ is such that such hierarchies have not really caught up with evaluation practice as desired in Africa. Further what is strategized as development in terms of the 3-pillars of economic, political and economic, have not taken on the aspects of progress that characterize Africa overtime since the strategies are not inclusive of multidisciplinary, multi-context, behaviour science and cultural nature that define Africa context. The situation is made worse by established practices that is drawn from many fields with situations as established by strong hierarchical systems that depend on meritocracy that have not really worked in Africa. Most are grand strategies and not the community level micro-features. Methods and techniques of ‘monitoring’ and ‘evaluation’ are also based on these grand strategies in terms of goals, outcomes determined from established knowledge and situation that arise.

**Objective:** Specifically the presentation will aspire towards the following:

- To revisit the commonly held 3-pillar view and suggest a multi-sectoral model in which **psychological** components arise and affect the programme programmatically
- To determine why after so many years, grand strategies development and performance in Africa becomes an outcome of ‘**conflict**’ \_ or lack of sustainable development evidenced

**Narrative and Justification:** The author would like to address this issue looking into policies of “Climate Change”(CC). The author will do this by building a theory of change approach for the problem object\_CC through articulating underlying beliefs and assumptions that will guide SDGs in Africa and are believed to be critical in producing change, a process to be designed to build on existing knowledge base.

AS a result, the purpose and approach will be:

- To ameliorate or restructure and re-model, a process that calls for building a theory of change approach for the problem object, climate change to guide practice in Africa
- To champion emergence of SDGs and Evaluation systems and regimes in Africa that will enable a targeted discussion of African situation on an on-going basis
- To enable brainstorming on the best solutions to meet the agenda for SDG across departments, inter-university, across other VOPEs and build a knowledge base and document the process thru publications

The author is of the opinion that failure to create such a guide leads to continued malpractice. The author underscores the urgency in this matter. She further contends from experience that Africa has needs to empowerment that should lead to sustainable strategies, systems and M&E practice. That Africa's status of being on the relieving end should be reversed and for Africa and her partners to be enabled thus creating an agency of sorts. That this is crucial to the process but it has to be done through a *theory change*. The support of the international community will be crucial. Empirical research should be part of this process both south-south and north-south. University environment is best suited for this task it being an academic body of research.

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 093 In Pursuit of Contribution Versus Attribution**

**O 270 - More Often than Not, Contribution and Attribution are Confused In Impact Evaluation. This Paper Splits These Elements Clearly**

S. Nyamhuno<sup>1</sup>

<sup>1</sup> USAID Southern Africa, Evaluation and learning, Johannesburg, South Africa

Impact evaluation is a challenging task to do due the rigorous nature of systems and processes that underpin the results. Many a times, those contracting the services of an evaluator are not themselves aware of what an impact evaluation entails. More strikingly, even evaluators themselves are not aware of what they are getting themselves into when they accept these contracts. In many cases contributions are mistakenly used in places of attributions.

Ascertaining causality to a particular intervention is a very difficult matter because one cannot without doubt directly attribute causality to an intervention because there are a lot of other possible causes for a change in a program. Attributions are often difficult to ascertain in complex evaluations because the results are prone to so many other influences, other than the intervention itself. It is evident, that change is seldom from a single factor, but a myriad of them. Long-term outcomes are particularly more prone to influence of so many externalities. An example is a 5 year HIV prevention program that results in HIV incidence declining by 5%. There is a possibility that apart from this program, print and electronic media campaigns from other programs contributed.

For one to be able to show an attribution, one has to clearly eliminate all external factors that could possibly influence the result. The only evaluation methodology that is capable of having an attribution is an experiment which mainly consists of randomized control trials (RCT) and quasi experiments. In experiments, an evaluator is able to have a very strong factual and counterfactual wherewith, one can able to remove all external factors that are likely to influence the results and then confidently attribute the results to an intervention. Surveys, natural experiments, correlation study and case studies can only result in contributions rather than attributions

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 093 In Pursuit of Contribution Versus Attribution**

## **O 271 - Addressing Unintended Effects in Development Aid Evaluations**

*A. Haslie<sup>1</sup>*

<sup>1</sup> Norwegian Agency for Development Cooperation Norad, Evaluation Department, Oslo, Norway

It has been known for decades that aid has unintended, sometimes negative effects in addition to the planned, positive results, and aid agencies have generally accepted and acknowledged that as a challenge. Despite this, unintended effects are often given less emphasis and are seldom systematically addressed in aid evaluations. Even many of the critical evaluations and general critique of aid often emphasize more the lack of positive (planned) results, than focusing on negative (or positive) unplanned effects.

The paper aims to explore how the evaluation methods and design can allow for a good measurement of unintended effects. Drawing on insights from research, both on development aid and on unintended consequences, and examples from development aid evaluations, the paper aims to contribute to a discussion of evaluation methods, design and criteria.

In evaluations of foreign aid, the quality standards of OECD-DAC are well-founded and recognized standards, and this states that evaluations should document unintended consequences of a development intervention. Despite this, a report commissioned by our department in 2014 show that this is not systematically addressed[1]. It may seem like the topic of unintended consequences undress an inherent dilemma in the design of evaluations; Evaluations are usually designed for measuring planned consequences of an intervention, how then can the evaluation design allow for a good measurement of unplanned/unintended consequences?

In the paper I will discuss how we can use insights from two bodies of research literature to address this more systematically in our evaluation of development aid. One bulk of research is the one concerning unintended consequences of development aid on both macro-, meso-, and micro-level. The other is sociological theories on unintended consequences, with Merton's typology as a foundation. In the paper I will also draw on examples from our own work with evaluations of Norwegian development aid.

Through the combination of insights from this literature, I argue that we can develop a methodology that allows for a practical, relevant and more systematic way of addressing unintended consequences in evaluations of development aid. In the paper, I will suggest an outline for what types of unintended consequences one could address in various types of development aid.

[1] «Unintended effects in evaluations of norwegian aid» May 2014, Norad

<https://www.norad.no/en/toolspublications/publications/2014/unintended-effects-in-evaluations-of-norwegian-aid/>

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 093 In Pursuit of Contribution Versus Attribution**

## **O 272 - Mixed Method Strategies for Strengthening Internal and External Validity: Combining QCA with QEM to evaluate household vulnerability in Uganda**

*C. Patterson<sup>1</sup>*

<sup>1</sup> *SoCha, Partner, Phuket, Thailand*

It is no secret that the use of experimental designs, such as Quasi-Experiment Methods (QEM), have risen to predominance in evaluation. Yet mixed-methodologists often critique QEM for its lack of explanatory power and substantive insight into how a project may have produced a positive impact. We bring this debate to a recent evaluation of donor efforts to reduce household vulnerability in Uganda. The evaluation design combines a difference-in-difference design QED to measure the project's impact with a Qualitative Comparative Analysis (QCA) approach to identify the many different ways the project may achieve success.

The combination of this approach yields a mixed method design strategy referred to as "lumping and splitting". For the sake of policy decision-making, it is oftentimes necessary to "lump" a wide array of outcomes and pathways into two groups (control and treatment) represented by summary statistics. This allows us to make clear, testable statements around program attribution and causal inference (aka internal validity). While useful, lumping seldom includes information on how the multitude of household pathways converged on these results across contexts, leaving us with little information on how to reproduce the results elsewhere. It is therefore necessary to "split" the control and treatment groups up into more refined subgroups to identify how local context and various aspects of a program combine and interact in different ways, allowing us to better understand how they can occur elsewhere (i.e. external validity). While also useful, splitting runs the risk of identifying pathways that may have been the result of false positives. As such, combining the strengths of both approaches compensates for their respective weaknesses.

We will briefly present the logic of the QEM and QCA, discuss the benefits of combining both, present findings to date and seek feedback on ways to improve this design's quality.

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 094 Evaluating Sustainable Development and 2030 Goals 2**

## **O 273 - National Evaluation Policies in Europe: Observations and Learning from the VNRs**

*K. El Saddik<sup>1</sup>*

<sup>1</sup> EVALSDGS, Vice-Chair, Ottawa, Canada

My contribution builds on the findings and observations of the latest two rounds of Voluntary Nations Reviews (VNRs) submitted to the High Level Political Forums by the European countries. VNRs intend to provide a stocktaking of the institutional, policy and implementation of the Agenda 2030 objectives. The paper will extend the work done by EvalSDGs-IIED and explore how the European countries have approached the follow up and review process at the national level. It will examine the various institutional enablers (focusing mostly on the national evaluation policy, systems and structures) that are gearing the national sustainable development agendas and map them using the Maturity tools suggested by the EvalSDGs Spotlight initiative.

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 094 Evaluating Sustainable Development and 2030 Goals 2**

**O 274 - Inclusive System Evaluation: Gender Equality, Environments and Marginalized Voices: A UN Women Guide for Empowering Voices in the SDG Era**

*E. Lewis<sup>1</sup>, S. Reddy<sup>2</sup>, A. Stephens<sup>3</sup>*

<sup>1</sup> University of Hull, Centre for Systems Studies, Kingston upon Hull, United Kingdom

<sup>2</sup> UN Women, New York, USA

<sup>3</sup> James Cook University, Cairns, Australia

This presentation introduces participants to a systemic thinking evaluation guidance, produced by UN Women. The ISE4GEMs is a new approach for the Sustainable Development Goals Era, which due to the many interrelated and interconnected SDGs, requires evaluators to think systemically, systematically and intersectionally. We introduce the GEMs framework – a framework for complex and systemic intersectional analysis which calls to attention culturally appropriate and ethical practices in evaluation planning, conduct, analysis and dissemination phases. The ISE4GEMs seeks to promote social transformation by understanding complex phenomena through a systemic approach and importantly, building evaluation capacity and every stage. The GEMs framework invokes an ethical imperative in the systemic methodological approach to the principles and practices to hear from different voices, values and forms of evidence to promote fairness, equity, accessibility and sustainability. This presentation will discuss both the theory and learned practice of its application with the UN and other global participants.

**Learning Objectives:**

By the end of this presentation, participants will be able to:

- List 1-2 activities that can be used to create a safe, inclusive, participatory environment for learning
- List qualities of systems thinker and an ISE4GEMs practitioner
- Describe key components of systems thinking and their relationship to conducting an ISE-4GEMs evaluation
- Describe different global cases studies that have used the ISE4GEMs methodology and how participants might use the evaluation plan of their own projects
- Have knowledge of the ISE4GEMs data collection and analysis phases, and where to find practitioner tools and guidance to support them to use this approach in these phases of their evaluation.

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 094 Evaluating Sustainable Development and 2030 Goals 2**

**O 275 - New Indicators for Sustainable Development Goals Targets; Technical Review of the Voices of the Hungry Project's Food Insecurity Experience Scale**

*N. Morrow*<sup>1</sup>

<sup>1</sup> Tulane University, Payson Program, new orleans, USA

In March 2016, the United Nations Statistical Commission's Interagency and Expert Group on SDG Indicators approved 230 indicators for the 167 targets of the 17 Sustainable Development Goals of Agenda 2030. At the end of 2017, 93 indicators have been assigned 'Tier I' status that implies that the indicator is conceptually clear, has an established methodology, and is regularly collected by many countries. The other indicators remain works in progress. Custodial agencies have been assigned to support member nations in collected the SDG monitoring data and leading the development of the indicators without clear methodological and conceptual guidance. This paper presents the findings of the technical review that was an input to the evaluation of the United Nation's Food and Agriculture Organizations Voices of the Hungry (VOH) project that successfully developed and tested a new SDG indicator called the Food Insecurity Experience Scale. This indicator measures the prevalence of food insecurity to monitor SDG target 2.1 2.1 "by 2030 end hunger and ensure access by all people, in particular the poor and people in vulnerable situations including infants, to safe, nutritious and sufficient food all year round". The review found that the VOH project team was able to leverage the capacity of a wide range of partners to create a new indicator that can be used for comparison to a global standard. The work on measurement modeling, in particular with expert input from the academic statistical measurement community, set this effort apart from other previous attempts at creating scales for food security measurement. Despite the technical achievement, interpretation with respect to some other traditional measures of malnutrition, poverty and food security of the new indicator posed challenges to use and uptake of the indicator for member nations wishing to monitor the SDGs. Restrictions on data sharing and a requirement for more consistent and targeted communications posed further challenges for establishment of the new indicator. Feedback is essential to sustainable development and the adaption and transformational aspects of resilience. New approaches and indicators for monitoring may be key to establishing the rich and diverse feedback at the foundation of resilience and sustainability. This findings from this paper may help inform other efforts at establishing rigorous methods and indicators for essential monitoring functions.

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 095 Evaluation In Multilateral Organisations

## O 276 - Evaluation as a Live Practice – in the Space Between Audit and Learning

*J. Gayfer<sup>1</sup>, D. Kabell<sup>2</sup>*

<sup>1</sup> IOD PARC, Director, Edinburgh, United Kingdom

<sup>2</sup> Kabell Konsulting, Consulting, Copenhagen, Denmark

Most evaluation policies state that evaluation is both for accountability and learning. Organisationally evaluation functions relate both to the audit function and to the knowledge management function, and can lean both ways, as primarily an accountability tool to help drive performance improvement and/ or tick boxes of compliance, or as a learning tool to spur curiosity, refocusing and/ or help shape new directions.

Evaluation – in its different forms – has the potential to change organisations, make them equipped for the future, more relevant, more effective and efficient, better places to work. But for this potential to be realized evaluation has to be forward leaning, a live practice where the use of evaluations emerges organically out of an evaluation system and process that engenders evaluative thinking and enquiry at the heart of the organisations drive for greater impact; engaging the right people, in the right way, at the right time and hence assisting the emergence of constituencies around its findings.

This paper applies this lens to the world of multilateral agencies, organisations that are – or are expected to be:

- Recognized global leaders in their field of intervention.
- Agile organisations that continuously track and adapt to global changes.
- Produce world-class products and world-changing results.
- Have state of the art business systems and processes.
- Attract first class staff from around the world.

It makes them interesting as models and inspiration for others who want to hold a mirror up to look at their own organisation to see how well it is doing, and where it might do better in terms of evolving its own evaluation system to one that is/ or remains fit for purpose in increasingly volatile and fast-moving times. A world where dealing with complexity is the norm.

We are basing our work on recent performance assessments of a large number of multilateral organisations for which the authors of the paper have been service providers. The assessments cover different types of agencies – normative, humanitarian, development, multilateral banks, thematic funds, inter alia – and draw out strengths and weaknesses and good and bad practice of evaluation functions situated within an overall view on the performance of the organisation. Our analysis and observations touches on issues linked to independence, organisation, quality, use, conditions for strong evaluation.

The specific objective of the paper is to examine where do these organisations currently position their evaluation system and practice in the span between accountability and learning? What wider organisational factors continue to influence this? How are evaluation systems changing? and what can we learn from this?

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 095 Evaluation In Multilateral Organisations

## O 277 - Integrating Gender in Evaluation in the Context of a Multilateral Development Bank

*M. Lomena-Gelis*<sup>1</sup>

<sup>1</sup> African Development Bank, Independent Evaluation Office, Abidjan, Côte d'Ivoire

The importance of including gender in development interventions has been widely recognized by all major Multilateral Development Banks (MDBs) at the highest strategic level and through specific gender policies, strategies and programs. Despite these normative advances, embedding gender considerations into designs, producing sufficient information about the distributional impacts among vulnerable groups and measuring the reduction of inequality gaps is still challenging both for operational departments and Independent Evaluation Offices of MDBs.

The gender-responsive evaluation literature recommends going beyond the usual practice of just capturing the sex-disaggregated numbers of beneficiaries or presenting average figures of the effects of interventions on women in general. Some promising ideas are related to undertaking a sound context analysis in order to craft gender indicators for each intervention. Others propose to look at the effects on different categories of women (intersectionality perspective) in order to capture the interaction between inequalities arising from race, class, gender, ability, geography and age. This calls for not considering women as a homogeneous group and include men in the evaluation of effects of interventions. It is also imperative to go beyond the usual measures of economic empowerment, encompassing issues such as reproductive health, legal and family codes, and tackling issues of voice.

These issues pose additional challenges to evaluators and require new collaborations and partnerships with operational departments, as well as with gender departments of MDBs, without jeopardizing the independence mandate of Evaluation Offices. Various Independent Evaluation Offices of MDBs have started collaborating within the Evaluation Cooperation Group (ECG) to take stock of the practices to integrate gender into evaluation (2016) and to produce hands-on reference documents to help integrate gender into different types of evaluation; the one for project-level evaluation was released in 2017. From its side, the African Development Bank has developed a normative work on Gender Equality since the approval of its gender policy in 2001 to the current Gender Strategy 2014–2018. For a number of years, the Bank has been screening the designs of its projects before approval to better integrate gender (one of the dimensions in the Readiness Review as part of ensuring Quality at Entry) and have produced several Country Gender Profiles to inform the strategic formulation of its partnership with regional member countries.

Building on various experiences from the African Development Bank's (AfDB) operational departments and the collaborative work with ECG, this paper will explore the emerging practice at the Bank's Independent Development Evaluation function (IDEV) to meaningfully include gender in different types of evaluation. It will finally place this reflection around the recent proposal at the AfDB of using a Gender Marker System at design. The paper will propose preliminary ideas about how ex-post independent evaluation could undertake gender impact assessments and summarize gender lessons learnt in the quest to inform the design of future development interventions which challenge existing social norms and the distribution of power and resources.

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 095 Evaluation In Multilateral Organisations**

**O 278 - Evaluation Function in the Spanish Cooperation:  
A Changes Proof System?**

*P. Caballero Partido<sup>1</sup>, S. Ulla Díez<sup>2</sup>*

<sup>1</sup> Ministry of Foreign Affairs and Cooperation of Spain, General Directorate of Development Sustainable Policies, Madrid, Spain

<sup>2</sup> Ministry of Foreign Affairs and Cooperation of Spain, General Directorate of Sustainable Development Policies, Madrid, Spain

The IV<sup>th</sup>. Master Plan of the Spanish Cooperation (2013–2016) established that the evaluation function in the Spanish Cooperation is a relevant activity to know the performance, products, results and impacts of our official development aid, focusing on generating useful knowledge and allowing the incorporation of learning. In addition, the evaluation is considered to contribute to the continuous management and accountability mechanisms improvement, to make informed decisions, ultimately, to a greater efficacy and quality of our interventions. This requires ensuring a more selective and strategic orientation of the evaluations, so as to optimize the available resources and increase the practical incidence of the conclusions and recommendations.

With this approach from 2013 to the present in the Spanish Cooperation system has advanced in the extension of the culture of evaluation among the actors, through the approval of an Evaluation Policy, the development of a biennial planning, the monitoring of the annual evaluation activity and the consolidation of a management response system, among others. However, since 2016, given the existing budgetary restrictions, a extended Master Plan, the provisional nature of the government for almost one year and the modification of hierarchical institutional structures within the department itself, have been able to introduce undesirable scale effects in the evaluation process of the central unit.

With this paper we intend to analyze these changes and context and what the consequences have been or not in the evaluation function, putting it to the test and what instruments or elements could be identified as key enablers for the continuity and maintenance of the progress achieved. All this from an eminently practical approach based on the experience of Evaluation of Development Policies and Knowledge Management Division of the Ministry of Foreign Affairs and Cooperation of Spain.

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 096 Leadership Perspectives

## O 279 - Performance Leadership: The Key to Strengthening Evaluation Systems and Program Outcomes

*J.S. Bayley*<sup>1</sup>

<sup>1</sup> EES, Department of Foreign Affairs and Trade, Yarralumla, Australia

(I request a time allocation of 30 minutes).

This proposal supports the conference theme 'Evaluation systems and approaches in turbulent times' by considering how we can:

- Build system leadership for evaluation policies, systems and practices, and
- Enhance evaluation capacities and culture to better undertake and use evaluations as part of a long term process of continuous improvement and creating public value.

Public officials around the world are searching for the one best performance system. For once they find this system, they will (they believe) have found a mechanism for improving their programs while strengthening accountability. And public officials everywhere are desperately trying to solve their performance problems. That is why they keep searching for the ultimate performance system. They won't find it. Why? Because the best performance system doesn't exist. Systems don't improve performance, leaders do.

The paradigm of evaluation as a tool to promote stronger reporting and accountability and hence drive program improvement has failed to meet expectations. It has failed because our theory of change is flawed. The evaluation profession would benefit from embracing a new paradigm of promoting and supporting performance leadership as the strategy for driving improved results.

To truly improve program performance within a context of constrained resources, active performance leadership is required, measurement and reporting alone are simply not enough. This presentation will provide an overview of the role and challenges for Performance Leaders in evaluation capacity building and achieving outcomes. It looks at how organisational leaders play a determining role in conceptualising and communicating an organisation's focus on continuous improvement. Senior leaders have a great many opportunities to strengthen their agency's focus on performance though:

- What they choose to attend to, where their times goes
- Their criteria for allocating resources
- Their reactions to problems and crisis
- The questions Performance Leaders ask at meetings
- Holding staff accountable for learning and continuous improvement
- Their distribution of rewards and status, as well as recruitment, selection, promotion, and retirement decisions.

Session participants will be encouraged to consider issues such as:

1. Does performance feedback inevitably lead to performance improvements?
2. Where do drivers for change come from?
3. Why are evaluation capacity building initiatives so difficult to implement? Why is it so easy for agencies to backslide after having made significant progress?
4. Is our theory of change for organisational performance improvement realistic?

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 096 Leadership Perspectives

## O 280 - Evaluating “Leadership”: Thinking Modular in an Organization-Wide Evaluation

*T. Wind*<sup>1</sup>

<sup>1</sup> International Development Research Centre, Senior Program Specialist- Evaluation, Ottawa- Ontario, Canada

This paper presents ideas to address two challenges evaluations often face. First, evaluations are pushed to provide useful feedback in short time-frames. Second, they are sometimes required to take broad aspirational strategies, policies and results and formulate evaluative frameworks that capture the things that matter most to both an organization and its stakeholders, including gender and social equity. As such, the paper addresses the first and second conference strands, particularly “evaluation design and management – current challenges and opportunities” and “rethinking evaluation methods, design and criteria”.

Evaluations often face daunting demands to provide quick feedback based on robust evidence about intangible but critical results – most often on relatively small budgets.

This presentation will share an experience of using “modular thinking” to address those demands for an evaluation at the International Development Research Centre. Building off the ideas of “agile” software development, this evaluation is producing modular, digestible outputs that build on one another, and build on other studies and evaluations going on within the organization. The modules are designed with the direction of advisory group that ensures the relevance of the different pieces and supports the use of the findings. The paper will reflect on the extent to which the evaluation has been “agile”, on what elements, and where we have failed to be agile.

The evaluation is assessing a donor's contribution to building leadership among organizations and individuals who undertake applied research to promote positive social, economic and environmental change in developing countries. As such, they contribute to evidence bases for sound development in turbulent times. Focusing an evaluation on their leadership capacities is a way to highlight the strategies and results that strengthens the resilience of these actors and the contributions they make.

Leadership is a “hard to measure” topic because both the meaning and evidence of the results are context-specific, there are many factors that influence whether leadership develops and is effective, and timeframes between interventions and results can be long.

The topic requires us to move beyond evaluating capacity strengthening, to exploring what leadership means, including perspectives within our organization, and among our grantees. Because the organizations and individuals we fund are across Africa, Asia and Latin America, we need to ensure that the way the evaluation defines leadership results are culturally and contextually relevant. In addition, while aiming to support leaders and leadership, my organization wishes to ensure we do not only support entrenched elites and entrench inequalities. So evaluation criteria needs to question how and to what extent results support gender equality and multiple axes of social inclusion.

The paper will present the ideas and frameworks that are guiding the study that could be useful to other organizations seeking to evaluate results that go beyond capacity building.

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 096 Leadership Perspectives**

## **O 281 - Examining Personal and Organisational Impacts of a Leadership Intervention on a Non-Governmental Organisation**

V. Horvath<sup>1</sup>, E. Harte<sup>1</sup>, Z. Middernacht<sup>1</sup>

<sup>1</sup> ODS Organisational Development Support, Monitoring and Evaluation, Brussels, Belgium

**Rationale:** This paper is concerned with understanding how best to capture personal and organisational impacts as a result of a leadership programme.

**Objectives sought:** The purpose of the evaluation was to gather qualitative information through interviews, focus groups, briefings, and a document review, to determine the overall impacts of a leadership intervention conducted in an international NGO. The main rationale for the leadership intervention was to enhance leadership for the health of the organisation, not simply to develop leadership qualities in an individual.

**Brief narrative and justification:** As evaluators, the assignment presented the challenges of understanding the contribution of the intervention to the individuals participating in the intervention; to the country offices implementing the leadership intervention to meet diverse internal challenges, and to the overall resilience of the global organisation. Insights from the process and the results can foster the conversation around strengthening NGOs to face organisational and political challenges, but also on the role that evaluations can play in advising and facilitating these processes.

Evaluators can draw a number of lessons from this experience, of which we outline a sample. Firstly, it is essential to understand the evaluation commissioner or funder's expectations for the evaluation, specifically what they hope to learn from the process and how they intend to use the evaluation results. Secondly, it is important to pay close attention to the cultural and organisational differences between an INGO's different country offices, and to tailor data collection tools to suit the context. Finally, the success of a leadership intervention at an individual or organisational level depends largely on the organisational set up, including resources to enable staff advancement, and time to practice and transfer the learning from said intervention, as well as the selection of individuals to participate in such an intervention.

Thursday, 4 October 2018  
15:00–16:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 097 Inclusive Evaluation Practice - Assumptions and Biases**

**O 282 - Decrypting Contextual Data to Reduce Bias: Case Study in Women's Rights Movement from Non-Western Lens**

*L. Boswell<sup>1</sup>*

<sup>1</sup> *International Consultant, Monitoring and Evaluation Expert, Temple Terrace, USA*

In light of the most recent #Metoo movement sweeping the United States, surveys, opinions, and even the more simplified question becomes laden with bias before the first answer can be given. Sensitivity in asking very intimate or traumatizing details already directs the data to be biased. We presume to hear the story as a counselor rather than an objective observer or evaluator. The presumption is also that the person or story relayed is factual with no other external factors.

These automated biases as a result of the wave of support for the movement and unwillingness to empirically question, (ask the difficult question that may appear insensitive or unsympathetic) does a huge disservice in our ability to decrypt the context of the data away from these emotional biases. This paper examines different approaches to sensitive topics and elevates the evaluation above journalistic roles and steers to a more empirical objective research and evaluator. In addition, this paper is also a follow up to the interpretation of women's rights in general with examples of statements made by western journalist and eastern journalists regarding Iraqi Women's Right between 2000 and 2012.

The journalists had inherent biases that were seeped in the contextual data. Recognizing hot button, incendiary, or inciteful words that we may consider normal is not always normal in another cultural context. As researchers and evaluators, one of our ethical standards is to reduce bias in our questions, surveys, and analysis. This is more difficult when the topic is fresh, and the survey results may not be what the movement may want to hear. Decryption of inflammatory or dramatic recounts down to a more reduced bias will make us more seasoned evaluators when we navigate and learn the effects of interpreting contextual data when conducting research.

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

## S 098 Resilience

**O 284 - Considering Policy Environment in Providing Policy Advice – Evaluation of Knowledge Sharing Program (KSP) as Foreign Aid**S.W. Lee<sup>1</sup>, J.E. Shin<sup>2</sup><sup>1</sup> Seoul National University, Graduate School of Public Administration, Seoul, Republic of Korea<sup>2</sup> Seoul National University- Graduate School of Public Administration, Center for International Development Evaluation, Seoul, Republic of Korea

With increasing global efforts in international development, foreign aid has evolved in its volume and forms/designs. Among them, policy advice (or knowledge sharing) is a method of providing assistance to developing countries that has received the spotlight of international society. The Republic of Korea's Knowledge Sharing Program (KSP) is a successful example of such aid, with increasing demand in many developing countries. However, actual adoption of the "advised" policy is not just decided from the KSP itself but also influenced by other environmental factors. This paper introduces empirical evidence into which policy environment factors are critical in the policy advice-to-adoption process using regression-type analysis of the database that was derived from developing KSP evaluation framework. With the analysis results, the paper suggests further policy implications for development programs and evaluation methods.

**Additional Justification:** The findings and argument of this paper are based on the evaluation framework development project for the Knowledge Sharing Program (KSP) in the Republic of Korea. KSP is a continuously expanding program designed to share Korea's development experience with developing countries all over the world. Since 2004, over 600 sub-projects were implemented, most of which take the form of policy research and advice. Through the evaluation project, the KSP sub-projects were systematically categorized; then a logic framework was developed for each type based on theory-based evaluation methods. This new evaluation framework is expected to provide useful evidence for improvement of the KSP.

Generally, impact evaluations of advisory development projects are considered to be complicated due to the difficulty in finding appropriate counterfactuals. In this sense, an introduction of the new KSP evaluation framework project and the methods used in the project would attract a vast amount of interest amongst the audience. The paper itself argues for a positive relationship between the democracy level (with the government function score as one of its composites) and successful policy adoption from policy advice, which is supported with empirically proven evidence using regression-type analysis on data from more than 600 KSP projects.

Overall, the findings of this paper calls attention to the importance of policy environment factors in implementing development projects. This is significant not only for the successful implementation of future development projects, but also for effective evaluation in measuring the pure impact of the development assistance project itself while excluding effects of external factors. With additional in-depth academic research and studies, our empirical findings have potential to lead to significant policy implications for international development programs, while at the same time contributing to developing effective evaluation methods.

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 098 Resilience

## O 285 - Building Resilience to Recurrent Refugee Crisis: an Evaluation of the Uganda Comprehensive Refugee Programme

*J. Kyewalabye*<sup>1</sup>

<sup>1</sup> Uganda Evaluation Association, Learning- Monitoring- Evaluation and Knowledge Management, Kampala, Uganda

**Rationale:** Uganda hosted the Solidarity Summit on refugees in June 2017 to raise resources. To help the situation, different countries and private sector players pledged financial and in kind support during the summit towards supporting the refugee effort. With so much money suddenly made available, for the refugees, the Uganda Evaluation Association (UEA) got concerned about how the money would be used to effectively benefit the refugees.

**Objectives:** As part of its capacity building program for its members, in 2018, UEA commissioned an evaluation of the Uganda Comprehensive Refugee Programme (UCRP). The objective of the evaluation was to study the effectiveness, efficiency, current challenges, and emerging lessons of the multisectoral, multipartner approaches; technical, operational and coordination in implementing the UCRP.

**Narrative:** Uganda has hosted refugees and asylum seekers from a number of countries for many years including the Democratic Republic of Congo, Sudan, Rwanda, Eritrea, and Burundi. It's currently Africa's largest refugee hosting state and has one of the most progressive policies worldwide for people fleeing conflict and persecution. According to UNHCR it is home to 1.4 million refugees, one million are said to come from war-torn South Sudan. With the South Sudan conflict in 2017, the numbers doubled and overwhelmed the country. The large numbers present problems that include; providing food and shelter, energy for cooking and light while preserving the environment, providing water, education, and addressing health needs, among others.

A combination of quantitative and qualitative approach to data collection and analysis was employed. A non-experimental outcome evaluation design was adopted for this evaluation to guide a qualitative study.

**Results:** The UCRP is important because of multiple emergencies unfolding in various parts of the world and the exemplary role the country has played in welcoming refugees seeking international protection. Multi-stakeholder, multi-government partnerships have been effective in managing the refugee problem in Uganda. However, with the increased influx of refugees from the neighbouring countries, there is no comprehensive refugee response framework in place to handle emergency situations. There is limited focus on the refugees and whether their basic needs for protection and assistance are being met, e.g., there were allegations of trafficking of minor girls and women to marry men who are not of their choice after paying a bribe to officials at the border crossing points. There are gaps in systems and in partner coordination, transparency, and accountability of all the funds and humanitarian aid and records. It is the Office of the Prime Minister (OPM) that records all refugees in Uganda and its figures are relied on to determine the scale of the humanitarian assistance. However, some partners are in dispute of the accuracy of the number of displaced persons and its staff and officials from the OPM are undertaking the biometric re-enrollment to certify the actual number of refugees. Managing refugees is a complex, multidisciplinary, and longterm commitment which requires significant investments of time, human and financial resources, support, and attention from all leaders at all levels and collaboration of host communities, national and international governments.

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 098 Resilience**

## **O 286 - Whose Resilience? Considering Gender and Other Systemic Power Dimensions when Assessing Resilience Capacities through Impact Evaluations**

*A. Pretari<sup>1</sup>, E. Febles<sup>2</sup>*

<sup>1</sup> *Oxfam Great Britain, Program Strategy and Impact Team, Oxford, United Kingdom*

<sup>2</sup> *Oxfam Intermom, Resilience and Livelihood Unit, Barcelona, Spain*

Individuals and groups experiencing different positions of power are affected by different climate-related risks or by the same risks differently, and may experience resilience building initiatives differently. As evaluation practitioners, we identify the need for greater exploration of how we measure, evidence and understand the diverse and often divergent resilience capacities held by those individuals and groups of individuals, and potentially different impacts of resilience programmes. This is because measurement approaches and evaluation designs are often focused on understanding a generalized experience, typically take the household as the unit of analysis, and are not as sensitive as they need to be to issues of power within and between groups. In addition, gender-sensitivity in climate change studies, and arguably in other fields, often comes from looking at differences in vulnerabilities of female-headed households and male-headed households (Djoudi (2016)). This paper presents methodologies that go beyond, by considering experiences of men and women, within the household.

Through four quantitative impact evaluations of Oxfam resilience programmes at different stages and in four different countries over the last 12 months, we trialled different gender-sensitive approaches to measuring resilience capacities at the household and individual levels, by combining different methods (Sense Maker, household surveys, individual surveys, focus group discussions at design stage, or at validation stage). We developed approaches of measurement of gender-sensitive and gender-differentiated resilience capacities. Gender being one dimension of systemic inequalities, we were sensitive to other power dynamics and their interaction with gender in the context of each of those impact evaluations, such as age, caste and religion, position within the family and wealth, that may affect people's experiences of resilience capacities.

This paper contributes to rethinking evaluation methods, by introducing designs of impact evaluations and measurement approaches to resilience that take into account systemic inequalities, gender in particular, and their intersections. This is an area that is receiving increased attention, as a better understanding and taking into account of systemic inequalities, gender in particular, is asked for to build resilient societies (LeMasson et al (2015), *Gender & Development* Volume 23, 2015 – Issue 3: Resilience, Sotelo Reyes (2017)). This paper makes a contribution to questions around measurement across different scales – individual and household levels (Bene (2018)) and gender-differentiated impacts and intersectionality in climate change studies (Djoudi et al (2016)). This paper also touches upon the literature on subjective measurements of resilience (Jones and Tanner (2015), Bene et al (2015)).

Thursday, 4 October 2018  
15:00–16:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 099 The Commercial Side of Evaluation

## O 287 - Consulting as a Second Act Career: Transition Issues

*G.V. Barrington*<sup>1</sup>

<sup>1</sup> *Barrington Research Group- Inc., President, Calgary, Canada*

There are many reasons why mature evaluators may be thinking about a new career as an independent consultant. Shifting political sands, downsizing, boredom, burnout – these can all make the work environment less rewarding or less attractive. What to do next? Is it time to retire or is retirement an outdated twentieth century phenomenon?

Today more than ever, older workers are considering extending their working life – dwindling/non-existent pensions, insufficient personal savings, high debt levels, rising health care costs, and a growing trajectory of useful years ahead. More compelling, though, is a growing realization that choice is a possibility, choice about where to invest your energy, how to develop your creativity, and how to be on the frontier of your own life.

Despite the many self-help books on career change, little work has been done on the personal characteristics, benefits, risks, and requirements needed for older workers to make a successful transition to independent consulting.

I am currently conducting a small study for an upcoming journal article, entitled, *Consulting After 50: Redirection and Reinvention for Career Evaluators*. It includes a series of in-depth, semi-structured interviews with approximately 10 independent consultants over age 50 who considered or made a transition from government/academia/non-profits to independent practice. The findings of this study will inform this presentation.

Topics will include concerns faced by older evaluators including personal suitability, barriers, and risks. Strategies to support career resilience will be also discussed including self-reflection, goal-setting, and planning. A skills inventory method will be shared. As time allows, a discussion will be held on other career transition issues.

#### **Learning objectives:**

- To consider personal suitability for a second-act consulting career
- To identify barriers to making a career change
- To identify ways to assess marketable skills.

Thursday, 4 October 2018  
15:00–16:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 099 The Commercial Side of Evaluation

## O 288 - Evaluators in the Gig Economy

*P. Alvarez*<sup>1</sup>

<sup>1</sup> UN Women, Independent Evaluation Office, New York, USA

The evaluation market is experiencing notable changes – both for evaluators working as staff in evaluation consulting firms or for independent evaluators. Within companies there is a trend to build ad hoc teams and hire capacity externally. While in the past, companies were stronger candidates in public bids on the basis of their collective capacity, team flexibility is now appreciated by contracting entities. There is less senior expertise within companies and independent consultants are sometimes requested to contribute or draft the bid proposal.

This paper seeks to explore market changes for evaluation firms and independent evaluators, as well as for organizations commissioning evaluations. How can we collectively improve the procurement processes, advertising, selection processes and communication? How can we add more value and transparency? Can we get rid of all practices that are not in sync with market realities? What is the meaning of all this in the particular case of feminist evaluators? How do they fit in teams, in proposals, in evaluation firms and as independent practitioners?

The evaluation job market is already fragmented. Evaluators are inundated with new approaches. Evaluation terms of reference request a broad number of approaches and evaluation proposals offer even more. There is a general call for innovative methodologies to capture impact. Evaluations get more and more complex as program implementation is no longer, if it ever was, a linear exercise.

Being a feminist evaluation firm is not easy to define and being a feminist evaluator may be a choice securing your position in a team but also distancing yourself from ever being considered as a mainstream evaluator.

Researching about the current evaluation job market and being branded as a feminist evaluator, conversations with independent evaluators and evaluation companies generated interesting discussions. Feminist evaluators responded with a variety of viewpoints. For some of them – those working with organizations whose primary focus is on gender equality and human rights – carving a niche as a feminist evaluator has been productive. Others found that the niche was too small and they risked being marginalized and never considered as “evaluators” as if the qualifier feminist supersedes anything it comes to accompany. However, that risk disappears if you combine your evaluation expertise with systems thinking, social justice or development evaluation.

Thursday, 4 October 2018  
15:00–16:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 099 The Commercial Side of Evaluation

## O 289 - When Power and Money meet Evaluation – Revisiting the “Matrix of Values”

M. Gutheil<sup>1</sup>, J. Hofman<sup>2</sup>

<sup>1</sup> Optimity Advisors, European Public Policy, London, United Kingdom

<sup>2</sup> RAND Europe, RAND Europe, Cambridge, United Kingdom

Significant levels of external support at every stage of the policy process is commissioned by the public sector to external independent entities. From policy reviews to impact assessments and evaluations, the means by which such work is commissioned and managed is usually subject to largely standardised rules and processes. External consultants / evaluators are used because public sector clients want them to solve specific problems, because they need an external impartial view, and sometimes external consultants are called upon to legitimise results or provide backing for controversial policy decisions.

While the consultancy market is booming, it might not have been subject to scrutiny from an ethical point of view. The client-contractor relationship is often based on a complex set of expectations and interactions, which can lead to dilemmas between commercial requirements on one side, and moral obligations related to research ethics on the other.

Potential for conflict lies in the possible pressure for external consultants / evaluators to deliver certain results. As discussed in related literature, with clients being the “funder of the service”[1], what is the influence on the evaluator’s value stance? What are the ethical norms if a client seeks changes to study deliverables which the evaluator does not feel are warranted by the research findings? How can consultants / evaluators overcome this pressure and related dilemmas that might impact on their professionalism, standards and independence? The paper will provide a review of existing ethics codes that guide consultant-client relationships.

Taking this discussion further, if evaluators are “powerless relative to their clients”[2], how shall they respond to – at times – unethical behaviour of their clients? With the wealth and breadth of evaluation research on anti-bullying programmes, how do evaluators and evaluation societies themselves cope with potential intimidation and harassment in their relations with clients and evaluation stakeholders? Are the “Matrix of Values” (Sheinfeld (1977)) and related value statements still fit for purpose and still valid today? These questions will be further discussed in the paper and possible answers will be presented.

In addition, the paper presentation provides an opportunity to engage with the evaluation community in order to exchange experiences, approaches and methods in evaluation independence and ethics, by drawing on the presenters’ and participants’ experiences, with the aim to:

- provide an opportunity for the audience to respond to key questions, as well as challenge and engage in a discussion with the speakers;
- generate a dynamic debate among the speakers and the audience to stimulate the sharing of various perspectives.

[1] Sheinfeld, S. N., & Lord, G. L. (1981). The ethics of evaluation researchers: An exploration of value choices. *Evaluation Review*, 5(3), 377–391.

[2] Sheinfeld, S. N., & Lord, G. L. (1981). The ethics of evaluation researchers: An exploration of value choices. *Evaluation Review*, 5(3), 377–391.

Thursday, 4 October 2018  
17:30–19:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 100 New Perspectives on Educational Evaluation

## O 290 - Adaptive and Robust Evaluative Reasoning in Dynamic Evaluation Contexts

*R. Pritchard*<sup>1</sup>

<sup>1</sup> Education Review Office, Central Region, Wellington, New Zealand

**Rationale:** Evaluative reasoning is critical to any effective evaluation process. It is fundamental to establishing the value of programmes and institutions we evaluate and has been described as “the ‘core’ of our discipline, the essential piece that makes or breaks evaluation” (Davidson, 2014, p. 32). It follows that in order to ensure evaluation that is rigorous and valid we need to understand how evaluative reasoning works in practice. This is especially the case in fast changing contexts, often fraught with competing interests and tensions, and where the methodology involves teams of evaluators. How evidence gets examined, justified and ‘reasoned’ to a satisfactory judgement of worth is essential for establishing validity and credibility.

**Objectives sought:** The focus of this research into evaluative reasoning was the synthesis process embedded in current review practice in New Zealand schools and early learning centres. It examines how evidence was used by review teams in synthesis, the iterative nature of the reasoning, what factors facilitated or hindered sound judgement and how decisions were determined on what to include in reaching robust and valid judgements on quality. It also explored how the findings align with concepts of evaluative reasoning in the evaluation literature. The overarching purpose was to generate understanding of how effective evaluative reasoning can be nurtured in complex and fast-changing contexts to inform and enhance evaluation judgements. The data base which informed these objectives were four case studies of school review practices with follow-up interviews, and an evaluation practitioner survey exploring reviewers’ experiences and perceptions of evaluative reasoning in the context of educational evaluation in New Zealand.

**Brief narrative and justification:** The synthesis process is an important element in the regular review of New Zealand schools and early learning services undertaken by The New Zealand Education Review Office (ERO). These evaluations are highly responsive to unique and dynamic contexts and promote stakeholder participation. The process involves small teams of evaluators working onsite within short timeframes alongside evaluands to design, gather, analyse and synthesise relevant information, including internal evaluation, to generate findings. Ensuring evaluative reasoning is robust, adaptive and transparent is a key challenge and focus for ERO.

Findings highlighted the importance of relational trust within teams, the use of frameworks for discussion, documentation of team deliberations, stakeholder involvement and evaluator competencies in promoting evaluative thinking. One major outcome has been to develop criteria for effective synthesis to make evaluative reasoning processes more visible in practice.

### References:

Davidson, E. J. (2014). How “beauty” can bring truth and justice to life. In J. C. Griffith and Montrose-Moorhead (Eds.). *New Directions for Evaluation*, 142, 31–4

Thursday, 4 October 2018  
17:30–19:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 100 New Perspectives on Educational Evaluation

## O 291 - How Does Evaluation Reshape Accountability Systems? Insights from a Case Study on Higher Education Reform in Italy

*R. Lumino*<sup>1</sup>, *D. Gambardella*<sup>2</sup>

<sup>1</sup> University of Naples Federico II, Department of Social Sciences, Naples, Italy

<sup>2</sup> University of Naples Federico II, Naples, Italy

This article reflects on the relationship between evaluation and accountability within complex organizations. A case study on higher education reform in Italy, assuming the institutionalization of a new evaluation system as its basic driver, is used to investigate how the interaction between dynamics of evaluation and competition results plays within the fundamental re-designing of the developing forms of accountability and distributions of power between state, professionals and citizens.

The Italian university system has undergone many changes and reforms during the last 20 years. One of the most important issue has been the introduction of new evaluation devices. Since the late 1990s, in such an area a large number of experiments have been carried out at national and university level. However, the results have not always lived up to expectations and the originally foreseen benefits have been continually postponed. Since the mid 2000s, the establishment of the National Agency for the Evaluation of the Higher Education System and Research, the Gelmini Reform of the Italian University (L.240/2010) and the related complex of regulations issued in these years have radically changed the evaluation of the Italian university. This is rooted within the austerity imperative (linked to the crisis and the huge public debt) and a widespread climate of distrust towards professionals.

Behind a bland rhetoric centred on universities autonomy, transparency and market flexibility, a process of re-centralisation and strengthening of state control on professionals and governance management takes place via evaluation requirements. In line with European Quality Assurance guidelines, self-evaluation and external evaluation tools have been introduced to measure institutional, organizational and individual performances next to the redesign of a formally independent evaluation authority.

The introduction of this new evaluation system has been increasingly associated with the strengthening of a centralized reward/penalty system which has raised several discussions and disputes among academics. The excess of simplification, the controversial definition of measurement instruments, – the paucity of bottom up participation processes in the overall definition of the evaluative devices are some of the most recurrent critical issues. Controversies have also been raised about the unintended and dysfunctional effects produced both on the functioning of organizations and on the conducts of individuals.

Such an emphasis risks of disabling our capacity to critically examine accountability systems, the way they are put in practice, the consequences they generate, not just on governance systems, but also on the accountability systems already in place. In a nutshell, our ability to learn from experience.

In such a framework, the paper aims at challenging the “self-evident truths” and the dominant conventional wisdom that embrace the National Evaluation System of Higher Education, by analyzing the vast array of technologies, techniques and procedures that constitutes the new evaluative techne of the Italian HE, showing how it reframes the developing forms of accountability and what intended and unintended effects it is producing.

Thursday, 4 October 2018  
17:30–19:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 100 New Perspectives on Educational Evaluation

## O 292 - Exploring the (Un)Intentional Consequences of Teacher Education Programming in Uganda, Africa

*R. Kane*<sup>1</sup>

<sup>1</sup> University of Ottawa, Faculty of Education, Ottawa, Canada

As classrooms within Canada become increasingly culturally diverse, there is a critical need for teachers who possess the skills and understandings to work across cultures (Smolcic & Katunich, 2017). Many teacher education programs in North America now include international service learning (ISL) opportunities as a part of their B.Ed. Degrees. These are generally short-term placements, where teacher candidates spend time teaching in classrooms in international contexts very different to those they are familiar with. The demand and popularity of ISL has increased rapidly over the past decade (Larsen & Searle, 2017), however, previous evaluations of such programs demonstrate mixed results.

While the benefits to students from ISL, is well documented in literature (e.g. valuable insight for teacher candidates on new forms of knowledge, personal growth etc.), the negative and unintended consequences of such programming on the local host communities, particularly student populations (e.g., ethical challenges and inequalities), is less understood. For example, Langdon and Agyeyomah (2014) suggest that communities that host volunteer placements and international practicums have become laboratories where individuals test their academic and career choices.

Due to the significant consequences of these unintended outcomes on local host communities, it is important that evaluators studying such programs focus on the unintended program outcomes, in addition to those outcomes mapped out in a logic model or theory of change. This is particularly important when evaluating programs implemented in developing nations and utilizing the resources and expertise of marginalized or vulnerable populations.

The objectives of this paper are to discuss the findings of a study that aimed to: (a) explore the extent to which ISP enhances cross-cultural awareness for teacher candidates and prepares them to interact more effectively with culturally and linguistically-diverse students in their home countries, and (b) understand the impact of ISL on local host communities, by focusing on both intended and unintended consequences of the programming. To guide our initial inquiry, we identified the following broad research question: To what extent are the benefits of ISL practicums reciprocated in host communities? To answer this question, we followed a multi-stage case study methodological framework. As part of this systematic process, we surveyed and interviewed teacher candidates and local host community members. We then qualitatively interpreted and thematically coded the material.

The findings of this study revealed that while teacher candidates report personal growth and value the ISL as enhancing their career opportunities there are unintended consequences on the children and youth from the host community including favouritism shown to different children and youth, presumptions of continued support, and frustrations which collectively lead to internal tensions within the host community. These findings illustrate the need to further explore the unintended outcomes of programming when conducting evaluations, particularly in the ifield of education and in the complex environment of developing nations with marginalized and vulnerable populations. Such mindfulness would ensure that evaluations are not only useful to program managers, but also help diverse groups of people ensure that programming efforts improve their lives and make all of our societies more resilient.

Thursday, 4 October 2018  
17:30–19:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 100 New Perspectives on Educational Evaluation

## O 293 - Evaluation of a Low Cost Mobile Device Initiative for First Year University Students

*L. Velloo<sup>1</sup>, S. Goodman<sup>1</sup>*

<sup>1</sup> University of Cape Town, Centre for Higher Education Development, Cape Town, South Africa

Personal mobile devices such as tablets and laptops, have become an integral part of the educational process for university students. These devices offer distinctive and expanded learning opportunities in contrast to traditional learning approaches. Research globally shows how advances in wireless and mobile technologies have resulted in personal mobile devices becoming pervasive, more convenient and less expensive. Students in developing countries like South Africa do not have access to these kinds of tools at the same rate as the global trends suggest. Mobile device ownership presents a substantial capital outlay beyond the means of many South African university students.

Access to technology assisted learning is a taken for granted in some contexts. This is not the case in South Africa where significant disparities exist. Within this context the researchers explored the prospects offered by low cost personal mobile devices in a higher education context. The aim of the evaluation was to assess the goodness of fit of low cost personal mobile devices for learning enhancement. The University of Cape Town's Personal Mobile Device programme was initiated to attempt to address some of these disparities. The programme aims to provide students with access to low cost tablets with the view of creating greater access to a flexible teaching and learning environment. This presentation reports on the formative evaluation conducted which sought to establish whether or not the programme's short-term, medium-term and long-term outcomes were achieved. The evaluation found that the programme played a crucial role in providing students with access to affordable mobile technology in the form of low cost tablets and has demonstrated enhancement of students' ability to learn in a flexible manner.

The evaluation provides insight for improved design of future low cost device programmes at higher education institutions in a South African context, such as development and implementation of monitoring plans, training and support for students and staff. Taking cognisance of the challenges facing first year South African students, such as poverty, crime, literacy challenges and barriers to access, the evaluation findings can contribute toward informing policy at an institutional and national level.

Thursday, 4 October 2018  
17:30–19:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 101 Promoting Use of Evaluation****O 294 - Recognizing and Acting on Several Levels of Evaluation Use: A Synthesis of Existing Models***B. Gauthier<sup>1</sup>, N. Kishchuk<sup>2</sup>*<sup>1</sup> Réseau Circum inc. / Circum Network Inc., Evaluation, Gatineau- Québec, Canada<sup>2</sup> Program Evaluation and Beyond Inc., Evaluation, Montreal- Quebec, Canada

For decades, reflection on evaluation use has been heavily shaped by Michael Quin Patton's "utilization-focussed" books. Yet, this is only one angle on evaluation use. This presentation will offer an integrated model of evidence use for policy making and program improvement that is directly applicable to evaluation work and program impact. We will compare, contrast and integrate the interpersonal factors emphasized by Michael Quin Patton (2012), the systemic factors identified by Gauthier (2016), and the situational factors addressed by the knowledge translation literature and the knowledge brokering literature (as well as Kirkhart, 2000). The paper will also offer an analysis of how these various independent variables may interplay and vary in strength of contribution depending on the type of use (Kusters et al., 2011). The comparison of these models will identify new avenues for evaluator action to increase the likelihood of their findings being used and the effectiveness of that use. This paper is related to Strand 2 "Rethinking Evaluation methods and methodologies" (especially "Data collection, analysis and reporting issues particularly in challenging contexts") and Strand 3 "Developing the field of Evaluation to promote resilience and action in critical times" (especially "Communicating, using and embedding evaluation").

Thursday, 4 October 2018  
17:30–19:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 101 Promoting Use of Evaluation

## O 295 - Evaluation Use and Resilience: Going Beyond Interventions

*D. Contandriopoulos<sup>1</sup>, D. Brousselle<sup>1</sup>*

<sup>1</sup> University of Victoria, School of Nursing, Victoria, Canada

As a discipline, evaluation rests on the idea that evaluative evidence has a central role in the betterment of societies. However, institutions worldwide struggle to identify, adopt, and implement in a timely manner the best – evidence-informed – policies and practices. This, in turn, impacts the efficiency, sustainability, and resilience of communities and societies.

A large body of scholarship, both within and beyond the field of evaluation, has focused on identifying best practices to strengthen the influence of scientific and evaluative evidence on decisions and policies. However, the complexity of policy-level knowledge transfer and exchange (KTE) processes has thwarted attempts to produce strong instrumental evidence on the “how-to”. Part of the problem is rooted in the fact that much of the literature focuses on discrete “interventions”. By discrete, we mean that the process through which evaluations are commissioned, conducted and their results communicated is usually conceived as a series of independent and autonomous events. By contrast, utilization processes take place in complex networks where actors are interdependent and where evidence use is neither linear nor discrete.

Recent developments in the fields of KTE, policy analysis, and structural network analysis suggest a redefinition of how evidence utilization processes should be conceptualized. More truthful conceptualizations should take into account that evaluative evidence is never considered alone but rather within broader information exchanges networks. This, in turn, shifts the analytical focus to the nature and functioning of those networks where heterogeneous actors (beyond researchers, civil servants, and managers), collectively take part in utilization processes which are both collective and systemic. By collective, we mean that they occur in systems with a high level of interdependency and interconnectedness among participants and where use depends on processes such as sense-making, coalition building, rhetoric, and persuasion. In the same way, we describe the processes as systemic, in that they involve a slowly evolving set of participants interacting over long periods. Discrete decisions or events are never the end of an identifiable process, as in step-based linear or circular models, but rather steps in a broader game.

Such a view prompts a shift in effect attribution. Most of the evaluation use literature is based on causal attribution models, in which intervention effectiveness is conceptualized as attributable to characteristics of the strategy, users, or producers. However, if the structure of interconnections between actors is indeed a core determinant of KTE effectiveness, those attribution models are inappropriate. To foster utilization use, what becomes crucial is understanding the utilization networks structures and its functioning.

The presentation will showcase how recent developments in the field of Social Network Analysis (SNA) and their application to the conceptualization of utilization processes can help foster our understanding of evaluation use. Such a perspective suggests we should shift our focus from discrete interventions to the structure of the interconnections between actors. In conclusion, the practical lessons of such a shift will be discussed and connected with examples of how more robust models of evidence utilization have a role to play to support resilient societies.

Thursday, 4 October 2018  
17:30–19:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 102 Public/Private Partnerships**

**O 298 - Ex-post Impact Evaluation for Public-Private Partnership Projects Using Multiple Stakeholder Perspectives**

*J. Oliveros<sup>1</sup>, A. Aibinu<sup>1</sup>*

<sup>1</sup> *The University of Melbourne, Faculty of Architecture- Building and Planning, Melbourne, Australia*

Public – private partnership project (PPP) is a procurement method to provide a public service using private funding and expertise. Whereas many mega infrastructure projects around the world have been built using this method, its impact has been a subject of scrutiny both by the media and in the academia. An ex-post evaluation of existing PPP projects can provide insight into the real impact of PPPs and provide a basis for improving future PPP developments and support strategic government agendas.

This article describes an ongoing research exploring the development of an ex-post impact evaluation tool for PPP projects from the perspective of multiple stakeholders. A multiple perspective approach is adopted due to the assumption that PPPs are long-term complex projects with many stakeholders that have legitimate and conflicting power and interests (public sector, private entity, users, and society/taxpayers). The proposed evaluation approach is called “Project Success Evaluation Pyramid Model” (PSEPM). It is essentially a model that identifies a big spectrum of success criteria from each of the involved stakeholders and then employs a set of judgement principles to assess the validity of all those criteria. For each stakeholder, a unique perspective is generated that contains the relevant information that supports the stakeholder judgement towards the project. It employs the Project Success concept from the project management discipline in which project success is defined as the “meeting of stakeholder expectations”.

This evaluation approach differs from a traditional impact evaluation by explicitly isolating and linking each criterion to a specific stakeholder, making them accountable for what they expect and perceived from the project. It responds to the existing conflicting views towards PPPs, which are enlarged by ideological positions. The aim of this evaluation is to make explicit and transparent all the possible views toward the project; and from that point, generate a judgement towards the real impact of the project. It judges a PPP project from a holistic view, which is considered one of the problems of existing evaluations; they usually address one specific view, such as economic efficiency, social impact, urban impact, political revenue, etc.

This approach is being tested in infrastructure projects; a PPP hospital in Australia and a toll road in Chile. The preliminary analysis of the tests show (1) the subjectivity of the stakeholder judgements, (2) the unreliable sources that some stakeholders employ to support their judgement, (3) the key conflicting issues that are taking place within a project, and (4) the big spectrum of arguments that are present in a public discussion.

This article contributes to the development of applied mechanisms to evaluate complex projects in which a “system” of stakeholders interacts. The theoretical contribution of this approach is (1) the systematic process involving all stakeholders (not only users and clients), and (2) the use of “stakeholder judgement principles”, which are “high-level” guidelines to assess the conflicting judgements of the different stakeholders.

Thursday, 4 October 2018  
17:30–19:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

## S 102 Public/Private Partnerships

### O 299 - How to Tackle Attribution in Public-Private Partnerships?

E. Sarmento<sup>1</sup>, M. Uzunkaya<sup>2</sup>

<sup>1</sup> Nova Business School of Management and Economics, Novafrica, Lisboa, Portugal

<sup>2</sup> Middle East Technical University, Business Administration, Ankara, Turkey

Have PPPs' positive impact translated into lower prices, more jobs, enhanced development and less poverty? Our presentation takes stock of the recent academic literature on the empirical impact of PPPs with an emphasis on impact evaluation methods for PPPs. It provides guidance on how to measure impacts with recourse to different tools: theory-based evaluation, logical frameworks, M&E systems and impact evaluation and highlights distinguishing features across sectors.

PPPs include additional complexities as compared to traditional procurement and are being increasingly asked to respond to the call to improve the economic and social value equation of investments.

The main challenge in evaluation of PPPs is the need to identify the incremental effects, if any, of choosing PPPs as an alternative provision method over a counterfactual, which would in most cases be traditional public procurement. The impact of public investment interventions (in particular infrastructure) is itself a topic of complex dynamics; identifying the incremental impact of PPPs over public procurement brings additional complexity.

There is a plethora of anecdotal evidence and case studies where the empirical impact of PPPs through attribution remains dubious. Empirical evidence is for the most part limited to studies of private sector participation in infrastructure investments, where PPPs effects are not in most cases disentangled. Whether PPPs have actually created value for money is still a matter of debate, thus making it critical to understand why and how value is generated.

The existing evidence shows that there is a positive and significant impact of private sector participation in access, quality of services, labor productivity, and reduction in technical losses. The magnitude of the impact varies by sector and size of the project and with the context, especially as it relates to the institutional and regulatory environment.

Beyond the literature view, we intend to provide a snapshot of the state of the art in this topic. We intend to conduct semi-structured interviews to international and national institutions to find out how the impact assessment of PPPs is currently being done and conceived.

It is also important to distinguish between design and implementation in PPP evaluation at this moment in time, as we are knowledgeable about how to design an evaluation of cross-sector partnerships, necessary in a world more interconnected where multi-stakeholder partnerships are developing, but on the practical side, there are very few cases of impact assessments through rigorous methods.

**The main contribution** of this study is to offer an up-dated perspective of how impact evaluation implementation is being perceived and conducted at the international and national level in the context of PPPs and provide a review of the tools available to design evaluations in different sectors as a guidance to future PPP evaluations.

Thursday, 4 October 2018  
17:30–19:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 103 The Practical Opportunities and Constraints of Evaluation

## O 300 - Evaluation Challenges and Opportunities within Corporate Workplace Sustainability Programming: Integration of SDG 8 Economic Growth and Decent Work

*D. Mutambara*<sup>1</sup>

<sup>1</sup> Strategic Agile Global Pvt Ltd Consultancy, Management, Harare, Zimbabwe

Based on her experience in evaluating corporate workplace sustainability programmes in the private sector companies based in South East Africa emphasizing on agriculture, the author shows the complexity, barriers, challenges and opportunities of evaluation in this sector. Private sector evaluation of workplace sustainability programming, should be designed with special emphasis on ontology and epistemology so as to understand context based on reality due to its complexity in terms of business goals. In addition methodology, correct target audience and the right evaluation questions is critical.

This presentation analyses the barriers, challenges and opportunities along the evaluation processes from inception level. It shall range from evaluation project design of the terms of reference, selection of the evaluation team, data collection and finally dissemination and use of findings. The presentation in addition, further describes ways in which the data yielded by the evaluations, develop data informed recommendations to enhance diverse decision making processes, and the right people who should be involved. Finally, the presentation discusses considerations and lessons learnt during evaluation of dynamic workplaces which involves community members within the business footprint, engaged through corporate social investment initiatives. Furthermore, it also presents diverse advantages and limitations of the different evaluation criteria and methods and their relevance according to the evaluation questions. The results has shown that:

Evaluation teams have used several methods based on their appropriateness to specific questions and the target audience addressed. However some evaluation teams have failed to design appropriate methods speaking to required needs on the ground which resulted to biased findings. In most instances the evaluation findings end up being rejected due to integrity and validity issues since recommendations given would not be data – driven. Data collection is subject to limitations depending on the methodology employed during evaluation process.

Although such mentioned methods are widely used in evaluation, they have their own diverse limitations, such as type of sampling applied which can be probability or non- probability, thereby a source of bias; and a lack of unambiguous validation procedures to test respondents. Furthermore type of data collection methods is critical in evaluation to avoid bias. If quantitative paradigm is applied as evaluation methodology, sometimes it does not capture much on reality based on context in this social sciences field which is different from natural sciences where it highly reliable due to its objectivity in nature. In addition, it use a static and rigid approach and so employs an inflexible process.

Furthermore presentation and dissemination of evaluation findings sometimes involves the wrong target and utilisation of findings do not serve its purpose, it can be shelved in cabinets and gather dust. In addition to non-utilisation of data driven recommendations, some of documented good practice is not disseminated for others to learn. The process of resolving the above discussed barriers, challenges and opportunities has identified key features of good evaluation design, implementation and use of information.

Thursday, 4 October 2018  
17:30–19:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 103 The Practical Opportunities and Constraints of Evaluation

## O 301 - Critical Analysis of the Challenges and Opportunities for the Evaluation in the Public Sector in Uganda

*J. Ojok*<sup>1</sup>

<sup>1</sup> Uganda Local Government Finance Commission, Administration, Kampala, Uganda

Evaluation is as old as the world itself and has moved side by side the journey of human civilization. Depending on how far one goes back, various roots of evaluation will appear. Genesis 1:31 says when God created the earth, the light in the darkness, the waters, the plants, the animals, and finally man, at the end of the fifth day, God saw everything he had made very good. He used criteria that enabled him make an assessment on whose findings he was able to make a fundamental decision. God's archangel asked God, how do you know that what you have created is 'very good'? What are your criteria? On what data do you base your judgment? What results were you expecting to attain? And aren't you a little close to making a fair and unbiased evaluation? "God thought about these questions and that day God's rest was greatly disturbed. On the eighth day God said, "Lucifer go to hell."

Monitoring and evaluation (M&E) has seen a steep climb within Africa—in terms of practice, profession and academic study. As a field of practice, specialized practitioners now exist and the demand for evaluation of policies, projects, programmes and interventions remains on the increase although legal and institutional frameworks for the practice are still weak.

Uganda like other African countries has been grappling with the challenge of increase efficiency and effectiveness in the delivery of services. The fundamental cause of these challenges has been attributed among others to weak M&E systems. To respond to this challenge government developed and implemented Public Sector M&E Strategy in the Ministries, Departments and Agencies with the view of enhancing accountability, transparency, management decision, organizational learning and promoting good governance. However, there has been a contention that a number of MDAs have not achieved the objectives of the M&E arrangement. It was established that accountability mechanisms for good governance and public service delivery are either non-existent or malfunctioned. It was proved that Ministries continued to post bad governance results. A study on decentralization in Uganda agreed with the fact that Public Sector M&E System is not geared towards understanding causality and attribution between stages of development change and that the quality of public service delivery is less than desirable and M&E system remained centered on compliance with government requirements and regulations rather than end-results of policy, program and project efforts.

This article examines: How M&E field evolved and what local and global forces have been behind this evolution? What precisely is the state of the M&E discipline in Ugandan Public Sector? What is the future of M&E in the Public Sector of Uganda? What are the challenges and opportunities in institutionalization of Public Sector M&E? Answers to these questions will provide useful insights into developing the field of Evaluation to promote resilience and action in the Public Sector Management.

Thursday, 4 October 2018  
17:30–19:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 103 The Practical Opportunities and Constraints of Evaluation

## O 302 - The Evaluator's Dilemma: Ethical Issues in Evaluations with Multiple Stakeholders with Different Agenda

*R. Dwivedi<sup>1</sup>, S. Mankad<sup>1</sup>, N. Sanwal<sup>2</sup>*

<sup>1</sup> Catalyst Management Services, Research, Bengaluru, India

<sup>2</sup> Catalyst Management Services, Research, Ratnagiri, India

The evaluation learning space is full of discussions on methods, tools, processes. Yet the involvement of different stakeholders and their differing agenda and politics is discussed less often; even though these aspects matter for the independence of evaluation, its ethics and outcomes.

The stakeholders in any evaluation study can be multiple and range from a few to many. These include the evaluation funder, its commissioner, the executing agency, the programme implementing agency. In some evaluations a quality assurance partner or even community representation may be involved in the mix. Each of these can be a single entity or multiple partners. The complexities in the evaluation multiply with greater number of agencies involved, with varied roles and stakes, and with the proposed use of the evaluation findings.

The authors of this paper have been involved in evaluations that range from the relatively simple to complex when it comes to the involvement and roles of partners and stakeholders. While some evaluations have had an evaluation commissioner who is also the programme funder, in others the implementation agency being evaluated has itself commissioned the evaluation. Programme donors in such cases may closely involved in the evaluation design and process or may be content with just receiving an evaluation report.

The stakes, and thereby expectations of different stakeholders also differ. While some evaluations search for the outcome or impact status and dwell in facts, some have a political or investigative undertones, and yet others may lean towards generation of learning for policy or scale.

The roles of the different participants in the evaluation also raise issues of ethics, independence of the evaluation, its use, usability and acceptance of findings by implementers. For instance, an implementer who is merely a respondent in the evaluation process and not involved in the design or analysis may not own the evaluation findings and question the interpretation. And yet, a more inclusive role where the implementer is involved in the data collection or the design risks positioning of the interpretations to suit the implementer.

Given these complexities, the questions arise: How much of independence should an evaluation claim? For whom is the evaluation needed? Who owns its findings?

This submission showcases five evaluations that the authors have been involved in with simple to complex stakeholder structures and roles. They share their dilemma in unpacking the varied expectations and understanding of stakeholders, and pinpointing primary stakeholder (Is it always the funder?). They instigate discussions around independence and, in fact, the sagacity in calling for total independence of evaluations. They present some techniques used to ensure the independence of findings while still maintaining the ethics and relevance of the exercise.

Thursday, 4 October 2018  
15:00–16:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

#### S 104 Building Capacity for Evaluation in Government

### O 303 - For Better Evaluation Capacity and a Stronger Evaluation Profession in Portugal: Challenges and Opportunities

*P.J. Teixeira<sup>1</sup>, A. Oliveira<sup>2</sup>*

<sup>1</sup> *Logframe - Consulting and Training, CEO, Lisboa, Portugal*

<sup>2</sup> *Portuguese Catholic University, Social Sciences, Lisbon, Portugal*

The authors will explore the challenges but also the opportunities and cultural transformative potential and value of investing in the creation of advanced evaluation training and university courses in a context where evaluation culture and practices are still generally lacking consistency – Portugal.

Starting with a brief analysis of the area of evaluation development trajectory in Portugal as it progressively gained more importance and recognition as an indispensable part of Policy, Programme and Project planning processes and how this increasing importance was greatly fuelled by the European Union funds.

This increased evaluation symbolic value and effective importance in Portugal promoted the creation of an evaluation market and the emergence of evaluation as a professional activity.

Will be characterised the evolution and current state of evaluation profession in Portugal and the stronger and weaker aspects of the evaluation market, giving special attention to the numerous reflection groups and communities of practice around evaluation that exist right now.

From this initial evaluation market assessment we can identify, in a structured way, a set of competencies that are necessary to the evaluators and that need to be taught and then put to the service of the evaluation processes in Portugal. In fact, one of the most important obstacles to the development of the evaluation profession and practices in Portugal have been the lack of quality evaluation specific university courses and professional training offerings. Until now, evaluation was only taught as part of other education and training offers and was never considered relevant to important enough to be the main focus of advanced, post grad or master offers in Portuguese universities.

This all changed in 2018 with the launch of the first post-grad evaluation centred course in Portugal at the Catholic University in Lisbon. The paper discusses the challenges to design, organise and launch this course when we have no previous experiences and have to literally develop something from zero. These challenges range from developing the course structure to find an experienced and qualified faculty.

As this is the first year of the post-grad offer, a critical analysis, made from different perspectives, of this first course will be presented and some clues for the future will be suggested in order to increase the quality of this specific offer, and also broader recommendations to the design of evaluation centred courses.

Thursday, 4 October 2018  
15:00–16:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 104 Building Capacity for Evaluation in Government

## O 304 - When Social Impact and Developmental Evaluation Meet in Japan – Professional Development of a New Breed of Evaluators

*K. Imata<sup>1</sup>, M. Nakatani<sup>2</sup>, N. Chiba<sup>1</sup>*

<sup>1</sup> CSO Network Japan, Tokyo, Japan

<sup>2</sup> Freelance evaluator, Tokyo, Japan

**Rationale:** “Well then, you don’t have to *unlearn* evaluation” was the comment we received from Michal Q Patton when we described that we would like to start a training program of developmental evaluator for SPOs (Social Purpose Organisations) in Japan. Compared to other OECD countries, Japan is a late-comer in mainstreaming program evaluation as means to promote accountability and learning in social impact programs. Only with the Government’s stipulation of the Policy Evaluations Act in 2001, program evaluation was made a common practice among central and local government bodies targeting policies, measures, administrative affairs, and executive agencies programs (policy evaluations). Nevertheless, for other actors such as non-profit organizations (NPOs) and community-based groups, program evaluation is still considered as something too cumbersome, too irrelevant and too expensive to apply to their programs.

Recently, a whole new wave of interest in program evaluations has emerged among NPOs and social businesses (we refer them as SPOs that denote non-profit or for-profit entities primarily organised to serve social purposes). One of the reasons is the Government’s new legislation to utilise Dormant Deposit Funds to fund organizations that promote public interest activities in which “social impact evaluation” is required to report on the grant activities. This is causing some uproar as many SPOs are not familiar with conducting program evaluation, let alone this particular evaluation approach. Moreover, there is a lot of confusion around “social impact evaluation” as it is still a half-baked amalgamation of program evaluation and social impact measurement.

**Objectives Sought and Brief Narrative:** In this paper, the authors will analyse the effort to introduce developmental evaluation (DE) to the Japanese SPO community. It will describe the training/professional development program started in the spring of 2017 whose goal is to deepen the understanding of DE in Japan and train key personnel of Japan’s SPOs to be equipped with DE approaches and capacities. The program, which has Michal Q Patton as the lead advisor and Kate McKegg (Kinnect Group) as the main trainer, is a two-year grant program to train 30+ SPO practitioners in DE, develop DE cases among Japan’s SPOs and disseminate the concepts of DE to the larger public.

In this paper the authors – who run this program – would like to elicit the findings of the following questions:

- How are Japanese SPOs finding advantages of DE compared to more traditional program evaluation approaches, as well as synergies between DE and new approaches in organisational theory and innovations in management?
- How is DE applicable or responsive to the call for “social impact,” especially in countries where the tide of impact investing has reached?
- What are the initial lessons learned from this Japanese experience to other countries where innovation is sought in using DE approaches to respond to the complexities of today’s world and seek sustainable and resilient societies.

Thursday, 4 October 2018  
15:00–16:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 104 Building Capacity for Evaluation in Government

## O 305 - A Survey on Organizational Readiness for Evaluation in Turkey

*S. Akgüngör<sup>1</sup>, Y. Gülcan<sup>1</sup>, Y. Kuştepe<sup>1</sup>, V. Erkan<sup>2</sup>, Ş.G. Çelik<sup>3</sup>*

<sup>1</sup> Dokuz Eylül University, Economics, İzmir, Turkey

<sup>2</sup> Ministry of Development, Planning, Ankara, Turkey

<sup>3</sup> Ministry of Development, Investment and project analysis, Ankara, Turkey

**Rationale:** Turkey lags behind other European countries in terms of the extent to which evaluation is considered as a fundamental component of governance and organizational decision making. Although there have been a range of projects and activities intended to put evaluation into use in Turkey, there is a need for improvement in evaluation capacity. Established in 2013, the Turkish Monitoring and Evaluation Society (TMES) aims to bring together professionals to practice and foster monitoring and evaluation culture across all institutions. As a registered voluntary organization for professional evaluation (VOPE) in the International Organization for Cooperation in Evaluation (IOCE), TMES aspires to develop a network of evaluators in Turkey and to share knowledge about evaluation with organizations that would potentially benefit from evaluation use. In recent workshops and consultations with stakeholders, TMES identified that a major problem is low demand for evaluation that creates weak capacity to do and use evaluation. Moreover, weak capacity restricts the possibilities for evaluation to be considered as an important learning tool in creating resilient societies.

**Objectives:** In designing evaluation capacity building (ECB) programs, it is important to know the baseline situation. A necessary step is to reveal the status of current understanding and practices of evaluation in Turkey. The objectives of the paper are to determine (i) the awareness of evaluation processes and practices and (ii) the level of readiness for evaluation in organizations in Turkey. The aim is to understand how evaluation is currently perceived, positioned and used in organizations in Turkey. Following the literature on organizational culture on evaluative inquiry (Cousins & Bourgeois, 2014; Mayne, 2008; Preskill and Torres, 2000), the study reports on the development and administration of a structured questionnaire with particular focus on various dimensions of organizational readiness for evaluation such as culture, leadership, resources and existing evaluation practices, if any. The survey sample includes representatives of government agencies and domestic and international NGOs in Turkey. The paper will present the survey findings and discuss options for ECB in Turkey.

**Narrative and justification:** This topic presents a baseline case study in a country where evaluation as a profession is newly emerging. The survey includes topics not only on organizational readiness for evaluation but also on organizational awareness on the use of evaluation. Results will be discussed in terms of their implications for ongoing inquiry into and practice of ECB. The paper will add-on to the findings of existing studies (such as Segone and Ocampo (2006) and Rugh and Segone (2013) on VOPEs by presenting the baseline demand for evaluation for a newly established VOPE. Moreover, the paper will discuss differences between government and non-government sector with regards to readiness and awareness for evaluation in a country setting where VOPE is newly developing.

Thursday, 4 October 2018  
17:30–19:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 105 Evaluating EU Policies and Programs**

**O 306 - Discovering Behavioural Patterns of Major Transport Projects: Lessons from an Ex-Post Evaluation**

*S. Vignetti<sup>1</sup>, C. Pancotti<sup>1</sup>, G. de Jong<sup>2</sup>*

<sup>1</sup> CSIL- Centre for Industrial Studies, Development and Evaluation Unit, Milan, Italy

<sup>2</sup> Significance, Director, The Hague, Netherlands

Project appraisal in transport is common practice for project selection and prioritisation, whereas ex-post evaluations are very seldom. Such analyses however can be useful to assess not only whether projects delivered the expected benefits, but especially to learn which mechanisms determined the observed outcomes and which determinants did affect the behaviour of the projects.

This paper presents the lessons drawn in the context of a recent ex-post evaluation of ten major transport infrastructure projects supported by the European Regional Development Fund (ERDF) and Cohesion Fund (CF) between 2000 and 2013. The developed conceptual framework heavily relied on a retrospective cost-benefit analysis which quantitatively assessed most of the direct effects generated by the projects, mainly in terms of reduction in travel costs. The study evaluates projects that are at least five years in operation. The fact that the CBA is carried out during the lifetime of the selected projects leads to a hybrid typology of CBA, sharing the features of both an ex-ante CBA and a pure ex-post (i.e. retrospective) CBA.

The most innovative aspect of the study however is that, building on the retrospective CBA, it tries to elaborate on the mechanisms explaining the observed performance and its key determinants. It does so by enriching the quantitative assessment of the CBA with qualitative analysis drawing on an extensive field work and interaction with key stakeholders, experts, and direct users. Field observations and the reconstruction of project histories enriched the purely socio-economic impact assessment with considerations linked to the political, sociological, institutional and cultural dimensions affecting the design and implementation of large infrastructure projects. Narratives and storytelling conveyed the main messages into ten case study reports.

On the basis of the literature on ex-post evaluation of major transport projects, we have identified five stylised patterns illustrating the typical behaviours of large transport projects, linking the observed outcomes (mainly in terms of cost and monetised benefits) with their determinants (in terms of governance, project selection process, forecasting capacities and other identified dimensions). Then, we tested the stylised patterns on the ten case studies, and, while generally confirming the validity of the classification, we could better fine tune it on the basis of real-case project histories. Relevant conclusions and lessons could be drawn on how to improve the design, selection and implementation of major transport projects.

Thursday, 4 October 2018  
17:30–19:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 105 Evaluating EU Policies and Programs

## O 307 - INTERREG Evaluations – The Challenge of Demonstrating Results and Value of Territorial Cooperation

S. Haarich<sup>1</sup>, M. Toptsidou<sup>2</sup>, G. Salvatori<sup>2</sup>

<sup>1</sup> Spatial Foresight GmbH, Germany, Troisdorf, Germany

<sup>2</sup> Spatial Foresight GmbH, Luxembourg, Heisdorf, Luxembourg

European Cohesion Funds, namely the ERDF with its INTERREG Programme, are nowadays an essential funding source for regional development in EU regions and beyond the EU borders. In this highly result-oriented structural funds environment, evaluation is gaining more and more importance. INTERREG programmes aim to promote cooperation across regions and Member States in the EU. For the European Territorial Cooperation programmes, also known as INTERREG, their evaluation is a necessary step during the programming period. According to Article 56 of the ESIF Common Provisions Regulation (1313/2013), “during the programming period, the managing authority shall ensure that evaluations, including evaluations to assess effectiveness, efficiency and impact, are carried out for each programme on the basis of the evaluation plan and that each evaluation is subject to appropriate follow-up in accordance with the Fund-specific rules.” The evaluation of the programmes shall allow to assess their progress as well as their impact on territorial cooperation and regional development.

Thus, each INTERREG programme defines its priorities and objectives, but also a number of indicators which would permit the continuous monitoring and eventually the evaluation of the different actions towards the targets set. Although the evaluation process may seem simple, it requires a continuous effort from the programme bodies to collect the necessary data and information and develop a monitoring system, but it also requires work from the evaluators to identify and coordinate the right methods to best develop an objective assessment of the impacts of the programmes. For INTERREG and its territorial cooperation goals, theory-based impact evaluation approaches are the most adequate. A mix of qualitative and quantitative methods usually is the most appropriate. Methods may include desk research of relevant programme documents, surveys directed to the different target groups, interviews with the programme bodies and other stakeholders, focus groups, case studies. These methods are tailor-made to each programme's objectives and needs and will allow to collect information from different perspectives. Specific evaluation challenges are related to the task to show impacts and added value on territorial cooperation, quality of life, economic development or administrative capacity and governance mechanisms.

Spatial Foresight in this Funding Period 2014–2020 is being involved in a wide range of INTERREG evaluations across Europe. Therefore, we have been able to develop a meta-evaluation of INTERREG programme evaluations offering insights into challenges and valuable approaches to evaluate and visualise results and value of territorial cooperation. We will show the pros and cons of different methods and present new approaches to show territorial impact of cooperation programmes. This is particularly interesting, as the new Cohesion Policy Framework of the EU post 2020 (and the future of INTERREG) will be decided soon. Cohesion Policy is challenged to provide better and more result-oriented outcomes and impacts to people and places. Our paper will offer a valuable contribution to this.

Thursday, 4 October 2018  
17:30–19:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 105 Evaluating EU Policies and Programs

## O 308 - Evaluating Job Quality in the EU Funds and Programmes that Support Smart, Sustainable, and Inclusive Growth

*J. Hofman*<sup>1</sup>

<sup>1</sup> RAND Europe, Home Affairs and Social Policy, Cambridge, United Kingdom

Following the 2008 global economic crisis, job creation has been a key focus of recovery. The EU budget funds several initiatives that support the EU's Europe 2020 strategy for growth and jobs in the current decade. Initiatives include the Horizon 2020 programme, which is managed directly by the European Commission (EC), and financial instruments such as the European Social Fund, where management is shared between the EC and Member States.

These programmes and instruments have different abilities and show various levels of performance in supporting and/or improving the quality of employment. There is some evidence on supporting job quality but the picture across the EU funds is patchy: with a few exceptions, there is almost no evidence on job quality at the EU level but often very rich data can be found at national and regional levels. These data, however, cannot be easily synthesized across the EU: the data focus on various (and different) aspects of job quality and draw on different definitions and methodologies. Yet, it seems that most commonly, the evidence addresses issues of job prospects and earnings.

The methodologies used for the assessments of job quality in the EU funds include survey methods and counterfactual analysis. Job quality indicators used in the EU funds focused on the dimensions that are easy to capture and directly measurable (e.g. type of contract, earnings). However, given the diversity of frameworks, indicators and indices available more widely, existing attempts to capture job quality in the EU funds can be considered as modest at best.

This contribution aims to assess and advise on how job quality can be measured in the EU funding framework by drawing on a scoping study that used document and literature reviews and stakeholder interviews.

The topic is likely to generate interest among those involved in the management, implementation and evaluation of EU funds and programmes – at the European, national and regional levels. The contribution is based on a scoping study that reviewed selected evaluations at the EU and national levels to identify methodologies, methods and metrics used to capture employment quality of jobs created (or supported) by various EU funds. The strengths and limitations of such approaches will be discussed in order to improve evaluation capabilities and provide more coherent and robust evidence in this space than it is currently the case. Given the challenges in current monitoring and evaluation arrangements for the EU funds and programmes, the contribution is expected to generate a heated debate on new ideas and improvements to the current systems. Job quality should not be overlooked within the EU funds and policy efforts need to consider it in conjunction with efforts to increase employment rates and investment. As such, this contribution seeks to advance the public interest through the promotion of equity.

Thursday, 4 October 2018  
17:30–19:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 105 Evaluating EU Policies and Programs

## O 309 - Evaluation Of The European Maritime And Fisheries Fund, Balancing Between Scarce Resources And Broad Thematic Scope

*A. Sanopoulos<sup>1</sup>, C. Hamza<sup>1</sup>, S. Mihalfy<sup>1</sup>, S. Schmuedderich<sup>2</sup>*

<sup>1</sup> DG Maritime Affairs and Fisheries, FAME Support Unit, Brussels, Belgium

<sup>2</sup> COFAD GmbH, n.a., Weilheim, Germany

The European Maritime and Fisheries Fund (EMFF) is rather small compared to the other European Structural and Investment Funds (ESIF). As all ESIF it is implemented through Operational Programmes (OP), one per Member State (MS). EMFF OPs are implemented under shared management (Commission and MS).

Regulation 1303/2013 defines in Art. 56 that OP Managing Authorities “shall ensure that evaluations, including evaluations to assess effectiveness, efficiency and impact, are carried out for each programme”.

The objective of the paper is to present the constraints faced by the EMFF OP when evaluating and to discuss on the procedural and methodological compromises necessary.

Compared to other ESIF OPs, EMFF OPs are relatively small (in average EUR 212 million EU Funds), however still have a broad thematic scope (fisheries, aquaculture, environmental monitoring, CLLD and local development of rural and coastal areas, processing and marketing and maritime surveillance).

Evaluating the EMFF OPs can be thus in challenge, especially concerning their specific weight in the sectors they target and the specificities of fisheries and maritime affairs.

The DG MARE monitoring and evaluation support unit (FAME SU) has developed a working paper on the evaluation of the EMFF OPs taking in account the constraints applying, the needs of the OPs, the size and resources of the managing authorities and the reporting obligations to the European Commission.

EMFF OPs have at their disposal a detailed and structured database on their operations, whose contents can be aggregated at different levels to provide different information to different users (policy makers, administrators, scientists).

MAs need to make “smart” use of their resources when evaluating. The choice of methods for evaluation should be adjusted to the available budget. This concerns the evaluation approach (theory-based or others) as well as the mix of qualitative and quantitative methods.

Thursday, 4 October 2018  
17:30–19:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 106 Evaluating Youth Culture Programs**

**O 310 - Evaluation Studies Supporting Complex Programs  
in the Field of Culture**

*A. Stawicki*<sup>1</sup>

<sup>1</sup> *University of Maria Curie-Skłodowska in Lublin, Philosophy and Sociology, Lublin, Poland*

The paper includes description of the experience gained during evaluation of the Very Young Culture program. Very Young Culture is the innovative program whose central aims are straighten organization operated in the field of cultural education and create trans sectoral network of entities cooperated in this field. The long-term goal was to shape young's people open attitudes, tolerance and the ability to participate critically in culture. The program's innovation lies in the way, how he was managed. The author of the program is prof. Marek Krajewski from University of Adam Mickiewicz in Poznań. The program had a nationwide range and was carried out at the regional level. The coordinator of the program was the National Center for Culture and the Ministry of Culture and National Heritage. The program operated on three levels – central, regional and local, and each of them provided for the implementation of an evaluation study. The evaluation research used the triangulation of both qualitative and quantitative methods to better capture the value of the program from the perspective of different stakeholders. The method of evaluation based on the theory of the program with the dominance of qualitative methods was used. The evaluation was continuous and was intended as a learning mechanism in the implementation of the program. The problem in the evaluation of the program was the immeasurability of long-term results. For example it was difficult to express the value of social and cultural capital, which are strengthened as a result of the program. In view of such long-term goals, an evaluation based on short-term variables, such as the self-assessment of competences acquired by program participants and the resulting cooperation network. The article will present the best solutions for measuring the results of the program realized in cultural field and the proposal to measure long-term results based on existing data banks from various social research.

Thursday, 4 October 2018  
17:30–19:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 106 Evaluating Youth Culture Programs**

**O 311 - Quasi-Experimental Impact Assessment: Evaluating Urban Centers of Culture, Art, Science and Sports in Brazil**

*F. Gonçalves Marques<sup>1</sup>, R. Antoniazzi<sup>1</sup>*

<sup>1</sup> *Plan-eval, Evaluation, São Paulo, Brazil*

In this article we examine the effects that attending the CUCA network has had on the lives of young people living in Fortaleza (Brazil). Through a quasi-experiment we carried out an impact assessment to verify the hypothesis that attending CUCA units positively affects young people's lives in the following aspects: family life; social-emotional skills; performance in studies; awareness of citizenship rights; preparation for the labor market; mitigation of illicit drug use; and awareness about how to avoid unwanted pregnancies and / or STDs. In addition, it was identified the possibility that the presence of the CUCA provokes an increase in the security sensation in the neighborhood where it is situated. To measure the above mentioned impacts, we use difference-in-difference regression models, in addition to Logit and Probit models. In this evaluation we face the challenge of developing a research design that uses the Control Group but in the Pretest.

This paper is part of an evaluation work carried out at the request of the Inter-American Development Bank.

Thursday, 4 October 2018  
17:30–19:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 106 Evaluating Youth Culture Programs

## O 312 - Building the Empirical Evidence of Youth Participatory and Collaborative Evaluation

*S. Heath<sup>1</sup>, K.A. Moreau<sup>1</sup>*

<sup>1</sup> University of Ottawa, Faculty of Education, Ottawa, Canada

Participatory and collaborative evaluation approaches require partnerships between trained program evaluators and program stakeholders (Cousins & Earl 1995) in the collection of information about program activities, to make judgments about programs, improve program effectiveness, and inform decision making (Patton, 1997). Due to the inclusive nature of participatory and collaborative evaluation approaches, organizations may view evaluation as beneficial, democratic, emancipatory, and empowering. Participatory and collaborative evaluation approaches have become popular among evaluators and program stakeholders as a way of developing organizational resiliency. In particular, through such approaches organizations learn to build their internal evaluation capacity and develop mechanisms for learning from their past work.

To provide guidance to evaluators and stakeholders, evaluation scholars (i.e., those conducting research on program evaluation) have conducted numerous research studies on the feasibility and effectiveness of using participatory and collaborative evaluation approaches in various contexts (e.g., Cousins & Chouinard, 2012). Yet, participatory and collaborative evaluation approaches are typically only involve adult stakeholders in evaluative processes. As a result, there remains a dearth of information regarding the feasibility and effectiveness of involving youth in participatory or collaborative evaluation approaches.

The objectives of this paper are to discuss the findings of a scoping review that aimed to: (a) summarize the extent to which participatory and collaborative approaches are used to engage youth in program evaluation, (b) highlight common participatory and collaborative evaluation approaches used to engage youth in program evaluation, and (c) identify research priorities to further investigate the use of participatory and collaborative evaluation approaches for engaging youth in program evaluation. To guide our initial inquiry, we identified the following broad research question: Based on the existing literature, what do we know about the use of participatory and collaborative approaches for engaging youth in program evaluation? To answer this question, we followed a multi-stage methodological framework. As part of this systematic process, we reviewed the published and grey program evaluation literature on participatory and collaborative approaches involving youth. We then qualitatively interpreted and thematically coded the material.

Despite the potential benefits (and challenges) of engaging youth in program evaluation, the findings of this review revealed that few documented program evaluations used participatory and collaborative approaches to engage youth in evaluation. These findings illustrate the need to further explore evaluation approaches used to evaluate youth serving programs, including the applicability and potential use of participatory and collaborative evaluation approaches. Such approaches may better engage youth involved in programs and services as well as provide insight to decision-makers and other stakeholders involved in serving youth populations.

This type of research on program evaluation will equip evaluators to engage youth in program evaluation and encourage the use of evaluation findings by stakeholders. To ensure that this paper is of interest to individuals working in various sectors, we also deliberated on the general benefits and challenges of using participatory and collaborative approaches for program evaluation with other vulnerable populations.

Thursday, 4 October 2018  
15:00–16:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 097 Inclusive Evaluation Practice - Assumptions and Biases**

## **O 313 - Capturing the Voices Of Marginalised People: Designing, Commissioning and Delivering Disability Inclusive Evaluations**

*A. Pollard*<sup>1</sup>

<sup>1</sup> *Government, Department for International Development DFID, London, United Kingdom*

**Rationale:** The Global Goals for Sustainable Development offer a historic opportunity to eradicate extreme poverty and ensure no one is left behind. To support the achievement of these goals, the global evaluation community needs to work with policy-makers to give voice to the world's most vulnerable and disadvantaged people. Evaluators also need to respond to the voices of those left furthest behind, such as people with disabilities, children, older people and those who face discrimination based on who they are or where they live. Although more than fifteen percent of the world's population has a disability (WHO, 2011), there are few resources that support evaluators and commissioners of evaluation to reflect on what constitutes good practice around disability inclusive evaluation.

This presentation aims to provide practical advice to support evaluation practitioners and commissioners to understand, design, deliver and manage evaluations in ways that are disability inclusive. It is based on lessons that were learned by evaluation consultants, a disability inclusion consultant and a commissioner of evaluations who worked together on a scoping study that aimed to understand how donors can support evaluation capacity development in ways that enable the voices of disabled people to be heard and reflected in country-led evaluation processes and systems. The team generated valuable insights that can support evaluators and commissioners to better understand disability and work in ways that meaningfully include people with disabilities, reflect their voices in evaluation.

**Objectives sought:**

The presentation will:

1. Facilitate a participative discussion about how to define disability and how this has implications for evaluation design
2. Reflect on how and why evaluators can and should work to consider disability from a rights-based perspective. This will include a discussion about why disability from a rights-based perspective remains poorly understood by many evaluators.
3. Provide practical suggestions about how evaluators and commissioners can work to design, commission and deliver disability inclusive evaluations
4. Discuss the relevance of inclusion to resilience: How and why are societies more resilient when they are inclusive?

The presentation will provide critical challenge and ask the audience to consider how the global evaluation community can work to ensure that disability is approached from an inclusion perspective; how commissioners can define and set expectations for evaluations to be inclusive and ensure that evaluation teams have appropriate technical expertise and training.

**Justification:** Disability inclusive evaluation is a highly relevant topic to the conference, as resilient societies are inclusive and the presentation will explore the relationship between resilience and inclusivity. The presentation is highly likely to advance inclusion and debate through discussion about diversity and equity within evaluation. Given that there are so few resources and little knowledge about how to engage people with disabilities in evaluation, the presentation will improve evaluators' capabilities and move the discussion beyond 'business as usual'.

Thursday, 4 October 2018  
15:00–16:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 097 Inclusive Evaluation Practice - Assumptions and Biases**

## **O 314 - Conducting Gender-Sensitive Needs Assessment. The Case of Caregivers Needs**

*J. Remr<sup>1</sup>*

<sup>1</sup> Charles University- Faculty of Arts, Department of Social Work, Prague, Czechia

The main goal of proposed paper is to summarize the experiences and provide recommendations concerning the methods of a needs assessment. It aims to improve evaluation capacity and provide persuasive arguments for re-considering the evaluation design. The paper offers useful tips and practical hints for evaluators concerning the evaluative research into gender-sensitive topics.

The case on which the arguments are built, comes from the field of social services, especially family care that is considered as a substantial component of social security system in many countries and it is a subject of interventions aiming to improve the lives of caregivers (i.e. those who are providing care to their elderly family members in their households) and care recipients as well. Satisfaction of caregivers' needs is essential for achieving the necessary quality, extent and adequacy of the care.

Findings from previous studies on caregivers' needs are usually biased due to the gender stereotype when caregivers are perceived as females. Such bias in perception has its roots in statistics showing that females represent the majority among caregivers. On the other hand, dominance of the "female agenda" within caregiving issues cannot justify the systematic neglect of specific needs that male caregivers have. Adjustment for such bias is essential because the needs assessment is often used as a source of the criteria for future evaluations.

Methodology of needs assessment, i.e. identification of the gap between current and optimal conditions for provided care was used. In contrast to other surveys focused on explicitly expressed wants of caregivers, the performed needs assessment enabled to identify the needs even of those caregivers who did not seek any help and who did not specified any requirements, preferences and wants. By using the needs assessment technique, it was possible to identify even the so-called unconscious needs. Such approach is crucial for gender-sensitive assessment because gender is a key discriminant variable when male caregivers express their preferences and wants significantly less than female ones.

Available results based on a series of approximately 60 in-depth interviews indicate that the combination of risk factors and needs differs between female and male caregivers. Presented experiences may help the other evaluators to propose improved design of their needs assessments. Presented experiences are relevant also for intervention-planners when they find-out the ways to control the reproduction of gender-based imbalance of the effects of social interventions.

Thursday, 4 October 2018  
17:30–19:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 108 Using Existing Evidence for Learning**

**O 315 - Multiple Outcomes of Building a Knowledge Management System to Manage Reflexive Impact Evaluations for Agricultural Research**

*G. Blundo-Canto<sup>1</sup>, A. Devaux-Spatarakis<sup>2</sup>, B. D'Agostino<sup>3</sup>, N. Mouret<sup>3</sup>, P. Martin<sup>4</sup>*

<sup>1</sup> CIRAD, UMR Innovation, Montpellier, France

<sup>2</sup> Quadrant Conseil, Quadrant Conseil, Bordeaux, France

<sup>3</sup> CIRAD, Direction générale déléguée à la recherche et à la stratégie, Montpellier, France

<sup>4</sup> CIRAD, UPR Aïda, Montpellier, France

The cross-case analysis of impact evaluations of research actions is a key tool to improve organizational evidence-based learning from both quantitative and qualitative data. Cirad, the French Agricultural Research Centre for International Development, developed a systemic *ex post* evaluation method, called ImpresS (Impact of Research in the South), to evaluate the impacts of agricultural research it carries out in developing countries. The method relies on four tools to describe an innovation process: 1) the innovation story, 2) stakeholder maps, 3) the impact pathway, and 4) the analysis of capacity strengthening situations. ImpresS was tested on 13 case studies, representing a diversity of innovation processes in which Cirad research was involved over the long term. The systematization of the information collected for the 13 evaluation reports provided a wealth of knowledge that was translated into a Knowledge Management System. This translation enabled different outcomes. First, cross-case analysis allowed the institution to better understand the nature and modalities through which its research activities lead to developmental impacts. Secondly, it supported the systematization of institutional memory and provided the basis for a formalized web-based interface that allows management, storage and export of *ex post* evaluations' knowledge. As a further outcome, the Knowledge Management System was upgraded to provide a graphic tool for *ex ante* scenario development. These outcomes of translating evaluation reports provide key insights into the value of knowledge management for institutional learning and support of research planning and implementation.

Thursday, 4 October 2018  
17:30–19:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 108 Using Existing Evidence for Learning**

**O 316 - The Missing Diagnosis and the Evidence Trap: A Tale of Inadequate Program Design**

*V. Dixon<sup>1</sup>, M. Gaarder<sup>2</sup>*

<sup>1</sup> *Think Global, Evaluation, Oslo, Norway*

<sup>2</sup> *3ie, Evaluation, Oslo, Norway*

To ensure that development interventions have an impact, making the right diagnosis is imperative.

The theoretical argument for why a well-planned development program that correctly diagnoses the main problems it is meant to address is more likely to succeed in addressing them, is generally accepted. Nevertheless, there is to our knowledge scant evidence to support this theoretical stance. Impact evaluations are a prime tool used to assess the effectiveness of development interventions on a variety of outcomes, to assess what “works” and what does not. When a large number of impact evaluations exist that look at similar interventions on similar outcomes, findings on what works and not may be synthesized in systematic reviews. However, impact evaluations and systematic reviews may not pick up on “misdiagnosis” or “missing diagnoses” of the problem to be solved.

What if we prescribed cash transfers to incentivize school attendance when the problem in a particular setting in fact was lack of quality teachers that led to children not attending school? Or what if we prescribed nutritional supplements when the problem was intestinal worms? Can we judge the treatment as ineffective when what happened was a mis-diagnosis and thus a maltreatment of the condition?

This paper will analyse a set of development programmes to test the correlation between appropriate diagnosis and the effect size of development outcomes, as measured in rigorous impact evaluations. Indicatively, we plan to do this across 30 programs covering several sectors. The findings will inform the debate on whether we are sometimes judging the effectiveness of the treatment based on a misdiagnosis by the ‘doctor’.

Linked to this is also the challenge of identifying and applying the appropriate medicine for the diagnosis. How do we know that a specific intervention will solve the identified problem? Development interventions are often designed to deliver the kind of interventions that the organization has specialised on delivering regardless of underlying causes to the problem (like, e.g. awareness raising of parents for increased school attendance), or interventions have been decided on by politicians or boards without appropriate preparatory diagnostic work. This paper discusses the potential implications of such practices based on our assessment of the correlation between appropriate diagnosis and development outcomes.

Thursday, 4 October 2018  
17:30–19:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 108 Using Existing Evidence for Learning**

**O 317 - What Can the Evaluation Community Learn from the Practice of Literature Reviews? Lessons Learnt from Eight Case Studies**

*V. Caby*<sup>1</sup>

<sup>1</sup> *Sciences Po Bordeaux, Centre Emile Durkheim CNRS-IEP-UB, BORDEAUX, France*

One major, although discrete, trend in the field of evaluation and expertise in France has been the development of a specific type of literature review oriented towards the evaluation of public policies: "the collective expertise method". Since the mid-90s, Ministries and organizations involved in the management of public problems and social controversies have ordered more and more of these evaluative literature reviews[1]. Yet, there has been very little effort to capitalize on them[2] – one reason being that public research institutes hold a monopoly over the collective expertise method.

Therefore, the aim of our paper is to present this evaluative literature review method, its strengths and weaknesses, with a focus on best practice that could be translated into the evaluation community. On a more theoretical level, another ambition is to confirm the hypothesis that the symbolic use of evaluation and expertise (their use as communication tools) requires the public openness reports (Weiss, 1979) as well as the scarcity of method (Beck, 1986).

This paper rests upon the comparative study of eight cases of implementation of the method. Cases range from obesity prevention strategies to strategies to reduce pain in animal farms. They are also diverse regarding the nature of the evaluation manager and the evaluator. All eight cases were studied using qualitative interviews with policymakers, evaluators and experts as well as documentary analysis.

We first describe the principles of the method, as well as its process. Using an extreme case approach, we underline several best practice regarding the selection of experts and literature, the management of the experts' group, and the dissemination of findings. Using policymakers' interviews, we demonstrate that the collective expertise method plays a "substantiating" as well as a "legitimizing" role within the policy process (Boswell, 2008). Such roles greatly rely on the method's narrative (a claim for interdisciplinarity and exhaustivity) and its process (the dissemination of reports).

[1] For an example: Vuillaume (2007) illustrates the growing use of this evaluative literature review within the Inter-Ministerial Mission for Combating Drugs and Addictive Behaviors.

[2] An exception is the special issue of the *Hermes* journal (2012) which includes several papers presenting the implementation of the collective expertise method within a number of public research institutes.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 109 Building Evaluation Capacity**

## **O 318 - The African Countries on a Path to Increase the Use of Evidence for Policymaking**

*M. Ramasobana*<sup>1</sup>

<sup>1</sup> Centre for Learning on Evaluation and Results Anglophone Africa (CLEAR-AA), Wits School of Governance, University of the Witwatersrand, Johannesburg, South Africa

**Background:** In Africa, monitoring and evaluation (M&E) short course training programmes have been perceived as a gateway to cultivate competent M&E practitioners and public servants. In addition, the training programmes seek to enable them to develop evidence-based approaches in the policy-making cycle, and in the prioritization of limited financial resources and results driven initiatives. Consequently, the role of the Centre for Learning on Evaluation and Results in Anglophone Africa (CLEAR-AA) as one of the prominent evaluation capacity building (ECB) pioneers within the African continent is significant. A total of 998 participants, with approximately fifty percent of the participants from various African countries, have undertaken M&E short courses delivered by CLEAR-AA during the period 2012–2017. The primary purpose of the proposed research study is to empirically measure how the short course training programmes delivered by CLEAR-AA contribute towards the improvement of participants' skills to execute their jobs and the broader evaluation practice.

**Objectives:** The primary purpose of the proposed research study is to evaluate the effectiveness of the short term course training programmes offered by CLEAR-AA in improving evaluation practice.

**Method:** The proposed research aims to collect data via in-depth interviews and conducting surveys with CLEAR-AA short course participants and preselected ECB pundits. In addition, an analysis of the CLEAR-AA evaluation forms repository will be conducted. The findings emerging from all the data sets will be analysed and compared with the literature on ECB.

**Results:** It is anticipated that the proposed research (including analysis) will provide information on/ into insights into participant's perceptions expectations and experiences pertaining to the effectiveness of M&E short course training programmes delivered by CLEAR-AA.

**Conclusion:** Based on the findings, conclusions on the effectiveness of the Monitoring and Evaluation short course training in Anglophone Africa will be provided.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 109 Building Evaluation Capacity**

**O 319 - Evaluating Monitoring, Evaluation, Accountability and Learning (MEAL) capacity: lessons learned from Save the Children**

*O. Kinda*<sup>1</sup>

<sup>1</sup> *Save the Children, Monitoring Evaluation Accountability and Learning MEAL, Dakar, Senegal*

This paper is primarily concerned with capacity building within civil society organisations (CSOs). Evaluating Capacity involves the design and implementation of teaching and learning strategies to help individuals, groups, and organizations, learn about what constitutes effective, useful, and professional evaluation practice. Baizerman, and Compton (2002), defines it as the intentional work to continuously create and sustain overall organizational processes that make quality evaluation and its uses routine".

Over the past years, Save the Children in Senegal has put strong emphasis on how to effectively strengthen the Monitoring, Evaluation, Accountability and Learning (MEAL) capacity of its implementing partners. Organizational Capacity Development (OCD) of partner organizations is one of the main strategies to strengthening them and making sure they are more resilient to challenging and changing contexts. As such, the capacity strengthening process is owned by the organizations (partners) and builds on their own priorities. OCD is then something that partners implement themselves based on their own priorities, not something that Save the Children does for them and this is expected to ultimately contribute to the overall objective and the sustainability of projects.

In the quest of taking their capacity to a next level and achieving equitable results, increased efforts with implementing partners have been made to ensure they have the requisite skills and competencies to improve programmes quality. This includes intensive, highly focused hands-on trainings to support indicator development and data-gathering mechanisms in selected sectors, and the establishment of peer-to-peer learning and exchange of experience. Despite efforts and increase attention, substantive issues remain. How should capacity be evaluated in order to meet quality standards? Which methods are necessary to evaluate MEAL capacity in a consistent and effective way? How powerful are methods and how can they best be influenced?

Based on recent experiences of Save the children in Senegal and West Africa regarding MEAL capacity with partners, this paper analyzes evaluation as a means to enhance capacity building. It highlights key challenges and benefits of evaluating capacity. It examines key methods used, and describes a variety of different tools and approaches used to plan, monitor and evaluate capacity building work. It provides an outline of current practice, based on interviews and consultations with key resource persons including evaluators. Finally, it highlights key areas for further discussion, and presents some conclusions drawn from previous studies. Lessons learned and good practices are addressed giving thus the opportunity to get meaningful insights on potential ways of addressing identified challenges.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 109 Building Evaluation Capacity

## O 320 - Applying the Peer Review Method to Civil Society Evaluations

*N. Gomes*<sup>1</sup>, *S. Beaujean*<sup>2</sup>

<sup>1</sup> Ministry of Foreign and European Affairs, Directorate for Development Cooperation, Luxembourg, Luxembourg

<sup>2</sup> ARTEMIS Information Management SA, Department of International Cooperation, Luxembourg, Luxembourg

When we speak about peer reviews, we think of examples such as the OECD/DAC peer review mechanism between its Member States or the peer review system of scientific journal articles by experts from the same field. Peer reviews are often associated with very technical and high-level exercises.

However, the peer review method can prove to be highly useful and relevant in different contexts. The Luxembourg Ministry of Foreign and European Affairs, Directorate for Development Cooperation, together with the West African NGO ENDA-Santé and the consultancy firm ARTEMIS Information Management S.A. have recently launched a pilot exercise of applying the peer review method to evaluating a regional civil society programme on HIV/AIDS in nine countries across Western Africa.

The programme, called "Frontières et vulnérabilités – FEVE 3", works mainly with key vulnerable populations, such as sex workers, men having sex with men, drug users or transgender people. These populations are hard to reach and can feel threatened by "outsiders", thus requiring a very sensitive and tailored approach. External evaluations of the programme have proven difficult in the past, as they were not always able to reach the populations and get the necessary data. As a result, ownership and implementation of recommendations have mostly failed.

In consequence, it was decided to pilot a new and innovative approach to make evaluations more efficient and ensure stronger ownership and implementation of the findings and recommendations. Essentially, the mid-term evaluation of the programme is going to be a country and peer-led evaluation process with a strong focus on learning. ARTEMIS was commissioned to assist in developing the multimedia toolkit and train peers to conduct evaluations and become trainers themselves.

In practice, each country is evaluated by two peers from other countries of the ENDA-Santé network and a member of the project's regional coordination team, respecting the principle of non-reciprocity. Having peers who know the issues faced by key populations carry out the evaluations ensures more openness and significantly strengthens ownership of the process, of the results and of the recommendations by civil society and government.

After the pilot phase in Senegal, Cabo Verde and Niger, the methodology will be applied across the remaining six countries of the programme and the peers who were trained during the first phase will become trainers themselves for the next round of peers. The Evaluation Department of the Luxembourg Ministry of Foreign and European Affairs, together with the regional coordination team of ENDA-Santé will accompany the process until the very end covering all nine programme countries.

The ultimate objective is to devise a methodology that can be taken to scale and applied across other civil society projects or any projects working with vulnerable and hard to reach populations and/or in challenging contexts where access is an issue. The focus in the pilot phase is just as much on training the peers and on learning as it is on the evaluation itself.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 110 Innovative Approaches To Impact Evaluation

## O 321 - Redesigning the Evaluation Architecture for More Resilient Evaluations; The Case of Integrated Care in the UK

*T. Ling<sup>1</sup>, A. Sutherland<sup>2</sup>*

<sup>1</sup> RAND Europe, Head of Evaluation, Cambridge, United Kingdom

<sup>2</sup> RAND Europe, Home Affairs and Social Policy, Cambridge, United Kingdom

Those delivering, commissioning and using evaluations often find themselves constrained in what they do by the way the architecture of the evaluation process is designed. This results in evaluations which may not address the important questions and may not build on previous evaluations to contribute to the knowledge of practitioners more widely. In this light, with funding and support from Age UK, the authors convened a series of workshops in the UK between November 2017 and April 2018 to explore these questions and this paper reports on these findings. We use the specific example of integrated care for older people but this in an effort to structure the discussion and findings are intended to be more generally applicable.

Evaluation systems involve many interconnected parts and processes including how the need is first identified, terms of reference set, bidding organised, contracts set, protocols amended, data collected and reports delivered. In a more interconnected and fast moving world this process has become dysfunctional. Using the example of integrated care for older people as a case study, we reviewed existing evaluations of this intervention to examine what research designs had been used and what problems had been encountered. This suggested that evaluation approaches were poorly suited to help deliver more resilient approaches to integrated care. This was followed by a series of workshops with researchers, commissioners of research and decision makers who used evaluations. We will discuss our findings suggesting that there are several avenues by which evaluation could improve, driven by more holistic and strategic thinking about the research and commissioning process and clearer ideas about what the end results of evaluations should lead towards.

Our aim is to open this important topic up for further discussion and to invite interested parties to take this debate forward.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 110 Innovative Approaches To Impact Evaluation

## O 322 - Designing Evaluations to Provide Evidence to Inform Action in New Settings

*E. Masset*<sup>1</sup>

<sup>1</sup> London School of Hygiene and Tropical Medicine, Center of Excellence for Development- Impact and Learning, London, United Kingdom

This paper is one of 14 inception papers produced by the DFID-funded, and newly established, Center of Excellence in Development Impact and Learning (CEDIL). The paper was produced by a multi-disciplinary team and led by colleagues at the London School of Hygiene and Tropical Medicine. CEDIL was established to address main gaps in evidence of effectiveness of development project and in impact evaluation methods in difficult areas such as fragile states and humanitarian assistance. In this paper we address the problem of 'external validity': how to extrapolate the results of impact evaluations from a context to another one.

Decision makers facing questions about how to address challenges in international development draw on evidence from other places, or from the past. In this paper we explore methods that evaluators can use to support the use of evidence to inform policy-making. We drew on consultations with experts and a scoping review spanning a range of disciplines and sectors. We identified concepts and approaches that can be used in the conception, design and analysis of evaluation studies to optimally provide evidence that can inform action in new settings. We described the concepts and organised the approaches under four headings: framing questions to test theories rather than interventions; process evaluations and mixed methods; leveraging heterogeneity; and using case studies.

From the concepts that we identified, a consensus emerged that the development and testing of a 'mid-range theory' was useful for learning more from evaluations, and that it resonated with other ways that interventions, mechanisms, and contexts have been conceptualised. From this position, we can learn more from evaluations when we ask questions that test underlying theories. Process evaluation and mixed methods approaches were summarised, identifying how these principles are guided by and can inform theory. Approaches for understanding how activities and contexts interact often take advantage of heterogeneity within and between evaluations.

We focused on individual evaluations, while recognising the critical role of synthesis in formulating policy. We reflected on how these principles support approaches to interpreting and synthesising evidence, finding a number of conceptual parallels and areas of synergy. Using examples, we reflected on the relevance of the principles for questions facing international development agencies.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 110 Innovative Approaches To Impact Evaluation

## O 323 - Lessons Learnt & Best Practices Managing Prospective Evaluations in a Constantly Changing Environment: Studies from Multi-Country Implementation in Gavi-Supported Countries

*L. Moreland<sup>1</sup>, E. Baguma<sup>1</sup>*

<sup>1</sup> *Gavi- the Vaccine Alliance, Monitoring and Evaluation, Geneva, Switzerland*

There is an increasing interest in, and application of, prospective evaluations to better enable concurrent, real-time evaluation of programme or policy implementation to improve the evidence-based decision making on design and implementation of key policy and programmes at the global and country level.

Over the past six years Gavi, the Vaccine Alliance has designed and managed multiple prospective evaluations, with the following principles defining its approach: flexibility, real time data and reporting, harmonisation and efficiency, country-driven, building capacity and utilisation focused. Multiyear, multi country prospective evaluations have been commissioned in a variety of countries with often complex operating environments, requiring creative and innovative management approaches. These evaluations have been designed in alignment with Gavi's Monitoring and Evaluation Framework, specifically its learning and accountability strategies, to facilitate the availability and use of robust information for learning, decision making and accountability for delivering on results.

The design and management of prospective evaluations requires innovative thinking and flexibility with respect to the evaluation design, including adaptive methods, sample selection and approaches; proactive stakeholder engagement; contract management; and timely dissemination of concise findings and actionable recommendations.

This paper will highlight the challenges and lessons learnt from Gavi's experience managing two prospective evaluations: Gavi Full Country Evaluations (FCE) (2013–2016) and Evaluation of Gavi's Cold Chain Equipment Optimisation Platform (CCEOP) (2018–2020).

- Full Country Evaluations in four countries (Bangladesh, Mozambique, Uganda and Zambia). The evaluation examined factors that influence successful immunisation programmes, obstacles to strengthening them, and the contribution of Gavi to improvements. Gavi's support for new and underused vaccines was assessed, along with cash-based support to countries. The prospective nature of the evaluation required careful consideration of real-time data collection, management and coordination with in-country researchers and Partners and flexibility of the adaptive evaluation approach.
- In 2017 Gavi commissioned an independent, multi-year prospective evaluation of Gavi's investments dedicated specifically towards cold chain equipment optimisation, at both a global and country level. The objective of the evaluation is to assess the relevance, effectiveness, efficiency, outcomes and sustainability of its CCEOP investment at a global, and country level (specifically Guinea, Kenya and Pakistan). The evaluation is taking place in parallel to CCEOP implementation, requiring careful consideration of planning and design to account for its prospective nature, specifically related to timing of baseline surveys, selection of intervention and non-intervention Districts/counties, ensuring a country-driven approach and managing constantly evolving country and global contexts.

Some of the key questions to be addressed will be:

- What are some of the key trade-offs that should be considered before embarking on a complex multi-year evaluation?
- What are some of the key considerations for managing dissemination, and use of findings, with prospective evaluations in a real-time, and diverse multi-stakeholder environment?

Reflections on the added complexity of conducting these evaluations in a prospective, real time manner in resilient countries, will also be shared to stimulate debate and discussion on how we, as an evaluation community, can continue to provide relevant, effective and sustainable support to resilient communities.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 111 Unpacking the Progress Toward the Sustainable Development Goals

## O 324 - Evaluation of Inter-linkages and Trade-offs Between the Sustainable Development Goals

*R.D. van den Berg*<sup>1,2</sup>

<sup>1</sup> King's College London, Department for International Development, Leidschendam, Netherlands

<sup>2</sup> IDEAS, President, Leidschendam, the Netherlands

The purpose of this presentation is to discuss innovative approaches for evaluation of the Sustainable Development Goals (SDGs), for further development and implementation in evaluations.

Several research programmes have started reporting on what is known about interlinkages between the SDGs. This includes synthesis of existing scientific evidence, as well as an analysis of gaps and frameworks for analysis. In the policy domain there is increasing interest in identifying where SDGs would reinforce each other, and where trade-offs should be considered. Given its role in providing evidence on progress towards the SDGs, as mandated in the United Nations Agenda 2030, evaluations of SDG policies and interventions will need to take interlinkages and trade-offs into account. The frameworks developed by the research programmes can and should be adopted and where necessary adapted in evaluations, as they provide a scientific perspective on linkages and trade-offs.

This presentation will focus on two frameworks for analysis that especially hold promise and will propose how these can be taken up in evaluations. First of all, these frameworks provide inspiration for the theory of change of why policies and interventions would contribute towards achieving a specific SDG. They provide a handy reference to potential interlinkages and trade-offs. Examples will be given of how a theory of change of an intervention could be enriched by using the frameworks.

Secondly, they indicate the nature of the interlinkages and thus make it possible to consider appropriate evaluation techniques. Again, examples will be given to illustrate possible evaluation tools and methods that will be "fit for purpose"

The frameworks that will be discussed are:

1. "A Guide to SDG Interactions: from Science to Implementation" of the International Council for Science, developed in collaboration with 8 universities and research institutes.
2. "Global Goals Mapping: the Environment-human Landscape" of the University of Sussex, UN Environment and WCMC

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 111 Unpacking the Progress Toward the Sustainable Development Goals**

## **O 325 - 10 Years of Monitoring Progress towards Sustainable Development in the EU – What Lessons can be Learned for the SDGs?**

*M. Hametner*<sup>1</sup>

<sup>1</sup> *Vienna University of Economics and Business, Institute for Managing Sustainability, Vienna, Austria*

The Agenda 2030 for Sustainable Development, adopted by the UN in September 2015, presents a new global policy framework, comprising 17 Sustainable Development Goals (SDGs) with 169 targets. Monitoring progress towards these 169 targets is particularly challenging, as exemplified by the still ongoing efforts of a dedicated UN working group, the IAEG-SDGs, to set up a fully-fledged indicator set for this purpose. The UN indicator framework, adopted in 2017, currently comprises over 400 different indicators; however, after more than 2 years of work, more than half (58%) of the required indicators are still not available, either due to a lacking calculation method or a lacking data collection from a majority of countries.

The main work on the SDG indicator framework still focuses on improving data availability; as a consequence, questions of how to comprehensively assess overall progress towards the SDGs and how to communicate the results of such an assessment to policy makers and the public have so far been more or less neglected.

Against this background, I would like to present insights and reflections from a decade of monitoring and evaluating progress towards sustainable development in the EU, based on my experiences as leading researcher in the preparation of the bi-annual Eurostat Monitoring Reports on Sustainable Development since 2008. In my presentation, I would like to highlight: (1) In which areas, and how fast, is the EU progressing towards the SDGs, and where is the EU lagging behind or moving in the wrong direction?; (2) How do different data sources and evaluation approaches affect the results communicated to policy makers and the public?; and (3) What have been the challenges of the EU's approach towards monitoring and evaluating progress towards SD over the past 10 years, and how may they be overcome?

In the frame of the EES conference 2018, the latter two points are of particular relevance, having in mind that evaluating progress towards the SDGs will necessarily have to build on a sound data basis. As such, lack of data, "problematic" (e.g. unreliable) data sources, incorrect interpretation of data, and simple calculation and aggregation errors by data providers are potential threats for conducting sound and reliable evaluations. Another challenge so far neglected in the discourse on monitoring and reporting progress towards the SDGs is grounded in the fact that all SDGs (and consequently all indicators) are considered equally important, thereby ignoring that some SDGs are not ends in themselves but means for achieving other SDGs. This discussion is different to the one on interlinkages between the 17 goals insofar as it is about the question whether sufficient emphasis is put on the important issues, which risk to get lost in the multitude of hundreds of different indicators.

In the proposed presentation I would like to highlight the most important challenges of monitoring progress towards sustainable development objectives at the EU level, point out solutions how these have been (or might be) overcome, and derive recommendations for monitoring and evaluating progress towards Agenda 2030 at the global level.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 112 Large-Scale Policy Evaluation

## O 326 - Socio-Economic Impacts from Innovation Scale-Up: A Framework for Evaluating Urban Innovation Programmes

*V. Valovirta*<sup>1</sup>

<sup>1</sup> VTT Technical Research Centre, Innovations- Economy- and Policy, Espoo, Finland

Cities are focal settings for innovation activities due to their high concentration of economic activity, skilled labour force, and diverse knowledge pool whose interactions are facilitated by geographic proximity. A variety of contemporary innovation policy programmes aim at stimulating urban innovation under the concepts of smart city, eco-city, sustainable cities and communities, and alike. It is argued that many grand societal and environmental challenges such as carbon emissions, traffic congestion, social integration and public health need to be solved in urban contexts thus calling for more innovation. Quite justifiably, urban innovation policy initiatives are increasingly becoming subject to evaluation and impact assessment.

Current models of programme evaluation do not sufficiently account for the complexity involved with how innovation activities generate social, economic, and environmental impacts in urban environments. Projects supported by innovation programmes typically focus on research, development, piloting and demonstration of new products, services, processes and practices. The immediate result of these innovation activities is new knowledge, novel or improved technology, a prototype, or the first instance of user tests. Connection of these knowledge outputs with socio-economic impacts on social wellbeing, sustainability, economic growth, and job creation is often distant and blurred. There is a gap between what impacts innovation programmes are expected to deliver and what they appear to be achieving on the short run.

This paper presents a conceptual framework for designing and evaluating innovation programmes with a focus on socio-economic impact generation. The approach builds on the programme theory and logical framework models which are commonly used to conceptualise policy interventions and their impacts. Inputs, activities, outputs and outcomes are distinguished as logical components of interventions and interconnected in causal impact chains. The challenge with using programme theory in evaluating innovation support programmes is that the connection of economic and social impacts with knowledge creation activities such as research and development projects remains un-articulated.

Our model extends the programme logic models by introducing innovation process as a complementary dimension. Building on innovation research we recognise that socio-economic impacts are generated only over time once innovative solutions – new products, services, processes or practices – become adopted and later diffused to a greater number of users. Instead of assuming one linear chain from inputs to outcomes we propose that innovation progresses through several stages. Successful innovation processes typically advance from research and idea generation to development, piloting, adoption and diffusion. Each stage is functionally distinct and later stages are conditional on the results of previous stages. Socio-economic impacts are ultimately created when innovation, that is a new or significantly improved product, service or process, becomes adopted in user practice generating net positive social and economic effects.

The paper demonstrates the use of the framework with an urban innovation programme case from Finland. Thereby we highlight critical steps to be taken into account when designing innovation programmes aiming to generate tangible social, economic and environmental impacts. The paper concludes by discussing implications of the framework for evaluation methods and practice.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

## S 112 Large-Scale Policy Evaluation

### O 327 - Evaluation of EU Assistance to Ukraine: How to Ensure Gender Balance

*O. Krasovska*<sup>1</sup>

<sup>1</sup> *The Institute for Economic Research and Policy Consulting IER- Leading research fellow/ projects coordinator, Centre for Economic Studies CES-IER, Kyiv, Ukraine*

**Rationale:** Revolution of Dignity in 2014 proved that Ukrainian society was ready to become the part of the European community. EU welcomed this decision that caused significant increase in numbers of various support programs for Ukraine via technical assistance, budget support, macro financial assistance, financing of infrastructure projects. This support covers almost all areas of reforms: decentralisation, SME development, education, healthcare, etc. Recent studies of the Institute for Economic Research and Policy Consulting and other experts prove that without EU support Ukraine was not able to implement all these reforms. At the same time question of gender equality and gender balance of these support is still open. The other important question is how to make evaluations of this support gender-responsive.

**Objectives sought:** Mentioned above questions justify objectives of this paper, that are follows:

- to study how EU support to Ukraine takes into consideration gender equality and gender balance;
- to analyze gender sensitivity of the evaluations of the EU projects made by independent evaluators;
- to provide recommendations how to increase gender relevance of EU support to Ukraine.

**Brief narrative and justification:** Gender issues remain low on the agenda in Ukraine, and there is still some resistance to discussing them. Unfortunately, in Ukraine there is no clear understanding of what gender equality and mainstreaming really is.

The majority of EU projects implemented in Ukraine took gender issues on board, but to the different extent. Ensuring gender balance amongst project participants is obligatory part of almost all projects while not all projects addressed more substantive aspects of gender equality, such as role of stereotypes and provide special activities (services) for women.

Analysis of including gender issues into agenda of EU funded projects in Ukraine will be provided in this paper. Because of the large number of projects implemented in Ukraine that are hardly assessed we will focus on the projects dedicated to the local governance and democracy, decentralization and SME development. The selection of these fields is explained by the strong influence of stereotypes there (e.g. male predominance in leadership and senior management positions).

The analysis will be done through publicly open sources according to the following criteria:

- does the project collect gender disaggregated data;
- does the project provide special services for women and other vulnerable groups of population;
- have the project implementing body special position of gender advisor.

Another question is planned to be answered in the paper is gender sensitivity of the projects' evaluations made by independent evaluators. It's planned to be done through analysis of methods used by evaluation teams while evaluating the projects activities and relevance.

Based on analysis of EU funded projects implemented in Ukraine recommendations on increasing their gender relevance will be elaborated. Recommendations on methods to evaluate gender sensitivity will be provided as well. Implementation of these recommendations allows putting gender issue into agenda in Ukraine and improving evaluations towards gender sensitivity.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 058 Defining Objectives and Indicators****O 328 - Rethinking Evaluation Methods, Design and Criteria for Assessing Policy and Programming Coherence when Evaluating Diplomacy, Trade and International Assistance Interventions**G. Gurova<sup>1</sup>

<sup>1</sup> Canadian Evaluation Society, Global Affairs Canada, Director of Diplomacy, Trade and Corporate Evaluations, Ottawa, Canada

Canada, and in particular, Global Affairs Canada (GAC) has a long history of evaluating diplomacy, trade and development programs and initiatives. The merger of Canada's International Development Agency with the Department of Foreign Affairs and Trade in 2013, coupled with a new Departmental Results Framework in 2017, which integrates diplomacy, trade, development and other programs, required new evaluation approaches, methodologies and criteria for assessing the extent to which the department is successful in building a coherent and synergistic approach to its programming. The Diplomacy, Trade and Corporate Evaluation Division is currently piloting an evaluation methodology for assessing the extent to which diplomacy, trade, development and security programs run by various branches of Global Affairs Canada are aligned with the broader Government of Canada policy objectives and with the needs of the host countries. The methodology aims at evaluating the extent to which programming is integrated to demonstrate policy coherence and, conversely – the extent to which policy is informing program design and objectives. The purpose is to identify the synergies that are being leveraged or built in the implementation of individual program interventions. The paper will outline the specific approach that Canada is taking to evaluate the level of policy and programming coherence across the "aid-trade-diplomacy" divide in Sub-Saharan Africa. It will discuss the evaluation methodology, related performance measures and indicators, potential challenges and limitations of the different methods applied for the purpose. Ultimately, it will identify and provide considerations and recommendations for future programming in other geographic regions with evolving political, social, security and economic contexts. Joint submission by: Galina Gurova – Deputy Director Diplomacy, Trade and Corporate Evaluation, and Stephen Kester – Director Diplomacy, Trade and Corporate Evaluation at Global Affairs Canada

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 112 Large-Scale Policy Evaluation****O 329 - The Role of Evaluation Agency for The Indonesian Government Policy in Water Access***D.R.A. Dewi<sup>1</sup>, M. Heldeweg<sup>1</sup>, K. Lulofs<sup>1</sup>*<sup>1</sup> *University of Twente, Department Governance & Technology for Sustainability CSTM, Enschede, Netherlands*

Today, the main challenge in water access policy is to establish an integrated, multi-stakeholder approach across national, regional and local levels in Indonesia. On the national level, the researcher defines four actors with four different roles relevant to achieving the current policy objectives, such as access to manageable drinking water in Indonesia. The Indonesian Government has set a target of achieving 100% of drinking water access Indonesia; as included in the Presidential Regulation No.2/2015 on the National Medium-Term Development Plan 2015–2019 (RPJMN 2015–2019). An evaluation agency has meanwhile been established by Indonesian Government: the Support Agency for Development and Implement of Drinking Water Supply System (BPPSPAM). BPPSPAM has a role in evaluating the performance of the drinking water supply agency from the Indonesian Stated-Owned Enterprises or Regional Owned Enterprises in improving their performance. Structurally, BPPSPAM is the agency which operates under authority of the Ministry of Public Work and Housing at the national level. This department carries responsibility for evaluating the regional drinking water supply agencies (the PDAMs). However, the regional drinking water supply agencies have no obligation to report to The Ministry of Public Work and Housing. Instead they report to the Regional Government under The Ministry of Home Affairs. This article's objective is to evaluate the governance structure of Indonesian water access policy activities by using the theoretical framework from Ostrom (2011), the Institutional Development Analysis Framework (IAD) that focuses on the interaction of actors (in this case the BPPSPAM as the evaluation agency and the other stakeholders) under the influence of rules-in-use and community attributes and biophysical conditions at the local level. The result of the analysis, a mapping of the relationships between these actors, clarifies the current policy ineffectiveness, particularly because of bureaucratic over crowdedness and missing links at and across various institutional levels. Furthermore, there are overlapping roles between different actors within horizontal and vertical relations, structuring and technical roles, across national, regional and local institutional levels. There is a gap between the practiced rules-in-use and prescriptive rules-in-form, particularly as a result of the community culture at the regional or local level. The BPPSPAM would be more effective as evaluation agency of the Indonesian drinking water supply system if it could operate as an independent agency equipped with the public authority to establish or terminate, and to intervene in the affairs of regional or local water agencies. To this end the IAD position, boundary, and authority rules concerning relations between the BPPSPAM and the relevant technical or administrative actors (particularly at national level) could be simplified to foster a more effective water access policy.

Friday, 5 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 113 Rethinking Evaluation Models for Challenging Times****O 330 - New Solutions for Europe ...and Beyond? The Role of Africa in Leading a Revision of the Realist Evaluation Lens***S. Chapman*<sup>1</sup><sup>1</sup> University of Cape Town, Management Studies, Cape Town, South Africa

Within our professional circles on the African continent, Realist Evaluation is regarded as something of a gold-star offering from the European Evaluation Community. On a continent where evaluation is often inadequately resourced or supported, realist evaluation offers a pragmatic way forward, when we don't have the luxury of starting every evaluation from scratch.

Now, the European Evaluation Society has called for new ideas in evaluation to address the state of crisis and turbulence faced in Europe *and beyond*. What does this mean for Africa and African evaluators – who are presumably drifting somewhere on the periphery of this process – in this so-called *beyond*. And what, in particular, does this mean for Europe's gold-star offering to African development evaluation – the realist evaluation lens? Will realist evaluation need to be transformed as a new offering *par excellence* to the African content – and if so, how can we as African evaluators lead (rather than simply follow) this process?

In this paper, we reflect on these challenges in light of our recent experience of trying to apply a "realist lens" to a characteristically African problem in crisis: the failure of many pro-poor policies in South Africa to deliver meaningful changes within the agricultural sector more than two decades on from the abolishing of apartheid. With this in mind, South Africa's Parliamentary Cabinet sought to address the crisis by calling on our team to deliver what they termed a "diagnostic evaluation" – a framework from which to move forward an integrated response to the findings of prior evaluations of policies aimed at supporting smallholder farmers. Because the objective of the diagnostic evaluation was to develop the basis (diagnostic) for a coherent set of evidence-based *mechanisms* that would work to support smallholders in different contexts, a realist lens appeared to us to be intuitively appealing.

Despite the intuitive appeal of the realist lens, we note in this paper how in the context of both a global and local South African political land crisis resulted in very real challenges around convincing policy makers as to the credibility of the so-called 'mechanism libraries' we were seeking to build. This process of re-framing evaluation findings in light of a political interpretative agenda, unless skillfully managed, can call to question the very assumption that evaluations can contain generalisable insights from across many different evaluations. Commenting as to what worked (and what didn't work) during this process, we build the case for how African evaluators could and should lead the discussion around the adaption of a Realist Evaluation discourse to future, more turbulent times in *both* Africa – and *beyond*.

Friday, 5 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 113 Rethinking Evaluation Models for Challenging Times**

**O 331 - The Rise of Realist Evaluations in Randomised Control Trials of Employment-related Interventions in the UK**

*J. Hofman*<sup>1</sup>

<sup>1</sup> RAND Europe, Home Affairs and Social Policy, Cambridge, United Kingdom

The UK government is undertaking a series of randomised control trials (RCT) of interventions that aim to support people with different health conditions into employment. The interventions are often based on the Individual Placement and Support – a fidelity model that has proven effective for people with severe mental health conditions. This and similar models are now tested with groups and populations that face different challenges into their integration with the labour market. The examples include: common mental health problems, musculoskeletal disorders, drug and alcohol dependence, to name a few.

Evaluations of such interventions comprise impact, process and economic elements to inform practice guides and implementation, if the models are found to be effective. Process evaluations integrate realist principles to examine and explain how the interventions have worked, for whom and under what conditions.

This contribution presents different approaches taken to address these questions in a number of realist evaluations that accompany RCT of employment-related interventions in the UK. The strengths and limitations of these approaches and challenges faced will be outlined to stimulate discussions and exchange of experience among evaluation practitioners. Such a debate is also likely to raise awareness of commissioners of evaluations on the realist approach.

The topic is likely to generate interest among evaluation practitioners and policy-makers in countries where RCT and realist evaluations of employment-related interventions are less common. The contribution is based on a series of realist evaluations that are carried out in the UK. The contribution is expected to generate a debate on improvements to the current approaches and alternative ideas. This will help evaluators to provide more useful insights about existing interventions and policy-makers to better understand and re-design mechanisms of labour market interventions.

Friday, 5 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 113 Rethinking Evaluation Models for Challenging Times**

**O 332 - Addressing Complexity in Programme Theory: Experiences in Applying the Actor-Based Change (ABC) Framework to International Development Programmes and Evaluations**

A. Koleros<sup>1</sup>

<sup>1</sup> Palladium, Research- Monitoring and Evaluation, London, United Kingdom

Programme evaluators regard a clearly articulated program theory as an essential first step for carrying out high-quality evaluations. At the same time, the need to tackle complexity to improve international development programming is widely accepted.

Despite a wide body of literature arguing that the success of programmes depends in large part on their ability to understand and engage with complexity, the development of useful programme theory for programmes intervening in complex systems still remains a challenge. Practitioners have to date not been able to effectively translate concepts from complexity science in a way which is coherent and consistent with the predominant frameworks used within international development practice. Programme theory, though often citing a recognition of complexity, continues to be largely articulated in the form of linear logic models.

Through our work designing and evaluating development programmes, we have developed the Actor-based Change (ABC) Framework as an attempt to address this gap. The approach uniquely focuses on the actors within a system as a pragmatic way of modelling how change happens in situations of complexity. Based on theoretical understandings of behavior change derived from behavioral science and new models of causal inference, the ABC Framework allows practitioners to effectively develop an evaluable program theory for interventions aimed at facilitating change in complex systems.

Since its development, the ABC Framework has successfully been employed on a diverse range of international development programmes for multiple purposes. These, including establishing monitoring and evaluation systems for adaptive programmes, designing theory-based evaluations of simple, complicated and complex interventions and programmes and conducting research into the effectiveness of interventions seeking to catalyse systemic change.

The ABC Framework was presented at the 2016 American Evaluation Association Annual Conference as a conceptual framework. This session builds the evidence base on its applicability and usefulness across a diverse range of international development programmes through a series of case studies.

From evaluating a security and justice strengthening programme in Nepal to establishing an adaptive MEL system for a transparency and accountability programme in Pakistan, this session will explore the strengths and weaknesses for this innovative approach to developing complexity-aware monitoring and evaluation systems and evaluation designs for evaluating complexity. It is hoped that session participants will gain a greater appreciation of the approach and how it can be adapted to their own work in developing programme theory for interventions and evaluations in complex and dynamic systems.

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 114 Professional Evaluation Organizations

## O 333 - The Roles and Sustainability Measures of Evaluation Networks and Associations in the Professionalization of YEEs in Developing Countries

*R. Michael<sup>1</sup>*

<sup>1</sup> *University of Kisubi, Faculty of Business and ICT, Kampala, Uganda*

Professional Associations are an epitome for collective bargaining, certification and accrediting of members. They are a hub for information sharing and a place where standards are set or revised about the professional integrity and conduct of the members.

This abstract is of a paper that focuses on the role of Evaluation networks and associations in professionalization and the sustainability of these networks in the professionalization of Young and Emerging Evaluators (YEEs) in the developing countries. The specific objectives for the paper are: to discuss the various roles that associations and networks contribute towards the professionalism of YEEs; to explain the various ways in which the networks and associations can be strengthened in order to enhance the YEEs' skills and lastly; to suggest the various sustainability measures for these networks and associations through which the young and emerging evaluators would grow professionally.

Among other roles, the paper focuses on the actual building of professionalism through networking, training and capacity building, encouraging of innovations and incorporating of new technology trends in Evaluation.

In Uganda, the Uganda Evaluation Association (UEA) has strengthened and sustained itself and its members through legal registration, new member recruitment from the Universities and higher institutions of learning especially those that have courses that specifically focus on either Project Planning and Management or Monitoring & Evaluation, social media networks, conferences and meetings.

As an individual, I have benefited directly from the Uganda Evaluation Association. This has been through information sharing. As a Young and Emerging Evaluator, I need the information at all levels if I have to develop in my profession. Information sharing could be in the aspects of the available opportunities to engage in for evaluation to develop experience, the new trends in evaluation, the available training and guidelines that direct the evaluation profession. This paper itself is a product of that information sharing experience where I have known about the call for proposals for the upcoming conference. The very awareness about webinars, EvalYouth, the EvalYouth International Mentoring Program for the YEEs are but some of the platforms I have encountered of which I can attribute to UEA.

This paper therefore elaborates on the roles of Evaluation Associations in general to YEEs in developing countries, how to strengthen these associations and suggest practical ways on how to sustain these associations such that YEEs in developing countries can, not only benefit from them, but be resilient and action oriented in critical times.

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 114 Professional Evaluation Organizations**

**O 334 - The Role of Voluntary Organizations for Professional Evaluations on Promoting Evaluation Culture: Experience and Challenges in Tanzania**

*F. Mwaijande*<sup>1</sup>

<sup>1</sup> Tanzania Evaluation Association- Mzumbe University, Administrative studies, Dar Es Salaam, United Republic of Tanzania

This paper examines the role of Voluntary Organizations for Professional Evaluations (VOPEs) in the developing a National Evaluation Culture. In the absence of a national Evaluation Policy in Tanzania, there is no evaluation culture. The evaluation culture is contextualized on the demand, use, and drawing lessons for future programming. This paper is developed on the premises that national evaluation policy determines evaluation culture. The objective of the paper is to describe the role played by Tanzanian Evaluation Association (TanEA) on promoting national evaluation culture. The source of data in this paper Organizational Capacity Assessment Tool (OCAT) for assessing the role played by VOPE. In addition to desk reviews of the available literature on the subject. The findings show that the established VOPES have roles for defining a national evaluation culture by fostering National Evaluation Policy. TanEA is pioneering the establishment of a national Evaluation Policy. The paper describes the process, procedures, and challenges in the process of developing evaluation culture. Lessons drawn would shade light to other VOPEs in the process of defining and promoting national evaluation policies and cultures for demand and use.

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 115 Monitoring and Evaluation Systems

## O 336 - Introducing a Monitoring & Evaluation System for the Russian Federation's International Development Assistance Program – A Look Behind the Curtain

*H.M. Boehmer*<sup>1</sup>

<sup>1</sup> *Duke University, Duke Center for International Development, Chevy Chase, USA*

The Russian Federation has long-standing economic and social ties and assistance programs for countries that used to be part of the Soviet Union. The nature of the assistance programs has changed significantly during the post-Soviet period and during Russia's first, and only, Presidency of the G8 in 2006. At that time, Russia committed to developing a Development Assistance program in line with the MDGs and international development practices. A Presidential Decree issued in June 2007 set out the first principles of foreign assistance, grounded in the MDGs and the Monterrey Consensus. Between the G8 Presidency and the 2009 global financial crisis, Russia's Official Development Assistance grew from about \$100 million to over \$700 million.

Despite the economic downturn and a reduction in the level of development assistance, the government issued a new Presidential Decree in 2010 enshrined the principles and goals of Russian development assistance program in law. While the new law still place Russian development assistance in the international context, there was a distinct shift toward national interests and an expectation that Monitoring and Evaluation would become a standard feature of all assistance programs.

Since 2015, there is a significant effort under way, largely driven by the Russian Ministry of Finance, to introduce a better understanding of the development outcomes that are being achieved through Russia's development assistance. There is now a strong interest in linking development assistance to the Sustainable Development Goals. However, given the decentralized nature of development assistance programs that are executed through numerous line Ministries, the process is politically highly sensitive.

This paper will briefly describe the development priorities and nature of Russia's assistance program, and subsequently focus on the Government's efforts to introduce the concepts of Monitoring and Evaluation through its bilateral aid programs, as well as through its influence in predominately Russian-owned regional development institutions, such as the Eurasian Fund for Stabilization and Development.

There are a number of interesting aspects emerging in political economy process of introducing M&E systems. While the bilateral introduction is fraught with political sensitivity, both within the Russian Government, as well as across long-standing bilateral relations, the multilateral process remains driven by Russia's significant regional influence. One interesting overlap is where former east-bloc countries have already joined the European Union, while central Asian countries are increasingly looking toward China or Europe, rather than Russian for their future development partnerships.

The primary learning experience from this paper is the challenge in adapting a system of learning and accountability for development outcomes to a culture of command and accounting in the aftermath, and expected future, economic and political turmoil.

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 115 Monitoring and Evaluation Systems****O 337 - No Pain No Gain: Controversies and Complexity in Modeling Mental Health Care Centers in the Amazon Region, Brazil**

É. Moraes<sup>1</sup>, E. Moreira dos Santos<sup>2</sup>, C. Rosario<sup>2</sup>, L.G. Martins<sup>3</sup>, L. Nascimento Pereira<sup>4</sup>, M.R. Oliveira Cardoso<sup>3</sup>, M. Batista Botelho Alves<sup>1</sup>, G. Cordeiro Pereira Cardoso<sup>2</sup>, S.C. Arias Bahia<sup>3</sup>, A.A. dos Santos Melo<sup>5</sup>

<sup>1</sup> Ministério da Saúde /Secretaria de Atenção à Saúde, Coordenação Geral da Política Nacional de Humanização CGPNH, Brasília, Brazil

<sup>2</sup> Escola Nacional de Saúde Pública/ FIOCRUZ, Laboratório de Avaliação de Situações Endêmicas Regionais LASER, Rio de Janeiro, Brazil

<sup>3</sup> Secretaria de Saúde do Estado do Pará, Coordenação Estadual de Humanização., Belém, Brazil

<sup>4</sup> Secretaria de Saúde do Estado do Pará, Coordenação Centro de Atenção Psicossocial Renascer/ Belém, Belém, Brazil

<sup>5</sup> Secretaria de Saúde do Estado do Pará, Coordenação Estadual de Humanização, Belém, Brazil

This work presents the modeling processes of the activities of a Centre for Psychosocial Care (CAPS) for building a monitoring system. The CAPS are part of the Psychosocial Care Network (RAPS in Portuguese) of Brazil's Unified Health System (SUS). Since its beginning, CAPS implementation design have as main functions to promote the deinstitutionalization (deployment of out-patient treatment services); the implementation of arrangements that complement and support family care; and the deconstruction of stigma among individuals with mental problems as dangerous to justify the need of seclusion. The case study is a CAPS of the State of Pará, in the Brazilian Amazon. Modeling is part of an implementation evaluation of the AcolheSUS project -a partnership between the General Coordination of the National Humanization Policy of SUS (CGPNH) and the Monitoring and Evaluation laboratory (LASER) of the National School of Public Health (ENSP/FIOCRUZ).

The AcolheSUS project is a collective construction on organizational development focused on improving working relationships aligned to the "methods" of the National Humanization Policy (NHP): inclusion of different actors; collective forums for conflict solutions; and promotion of networks. The ethnic composition of the Amazon region, due to the presence of several native population tribes and social inequality, poses a unique challenge for psychosocial attention in the SUS.

The reflective process of monitoring and evaluation (M&E) is aligned with the National Humanization Policy guidelines. It proposes M&E as a shared and co-managed activity committed to the reality of the health services and of its workers, assuming the inseparability between M&E capacity building and intervention (SANTOS FILHO, 2009). Additionally, the CAPS structure and linkage with its entitled territory is seen as socio-technical device (LATOURE, 2000) able to mobilize, share and produce knowledge and meaning. In this, the human factor is critical in the struggle against the hospital centered approach attention for individuals with mental health problems.

The goals of modeling CAPS functions were to clarify its theory of change and to identify guidelines for its monitoring system. Shared modeling occurred through four participatory face-to-face workshops, three virtual meetings, and participation in two users' meetings involving evaluators, managers and local implementers.

According to Davies (2005), controversial and complex processes are hardly modeled by linear diagrams. The CAPS developed model, involved embedded two chains (clinical psychiatric and psychological care) interwoven by “bridging” activities, such as users’ interactive meetings, case discussions and integrative conversation aligned to HNP guidelines. The linear modeling despite of embedded and hierarchical, failed to incorporate the relationships between CAPS and the community social support network. Several presumed chains of actions and effects are still invisible, overshadowed by controversies, such as: mental health care in its psychosocial dimension in the family and community; the recovery of the strategies developed by users to access alternative care; the changes in the representations of the mental suffering; and the recent intensification of pressures towards a resumption of the hospital centered model for mental care in Brazil. These last two a vital space of connections and inscriptions of resilience.

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 115 Monitoring and Evaluation Systems

## O 338 - Towards a General Theory of Monitoring and Evaluation: A Standard for the Scope of the Evaluation in Terms of Reference

*E.E. Goetsch*<sup>1</sup>

<sup>1</sup> University of the Witwatersrand, School of Law, Johannesburg, South Africa

**Rationale:** By mischief or mistake, implementation agencies can conceal embarrassments from external evaluators by tweaking the Terms of Reference. By restricting our ability to discover and disclose, the agency can avoid accounting to donors or beneficiaries any diversions of funds or defects in delivery. A standard for the scope of the Terms of Reference removes the risk of an accidental or deliberate cover-up, bringing transparency to projects, independence to the profession, comfort to the donors, results to the beneficiaries and rigour to the discipline.

**Objectives sought:** To propose a model Terms of Reference that satisfies role-players and stakeholders as accurate, complete, consistent, equal and precise when scoping the evaluation of any project in the international development aid sector.

**Brief narrative:** The author has occupied or advised almost all the positions in the international development aid sector. Especially, he has advised the managers of donors and implementation agencies in their day-to-day decisions. The model presented satisfies the occupational requirements of 10 areas of management in all fields of development and across the funding chain. It is accurate, insofar as it captures the key performance indicators in their correct positions. It is complete, insofar as it stretches from the recognition of the problem to the report of the solution and the 9 steps from funding to delivery. It is consistent, insofar as it applies equally to each and every project. It is universal, since it applies in every development field. It is precise, insofar as it treats the quantitative and qualitative elements mathematically. It is equal, since it is independent of gender, race or person. It is inter-sectional, since it recognises the perspectives and requirements of all stakeholders and role-players in the international development aid sector. It is practical, insofar as it was field-tested in resource-poor environments.

**Justification:** The model expresses a quality standard for a universal Terms of Reference. It has many benefits. It helps evaluations help people to improve their lives and make our societies more resilient. It reduces unpredictability and complexity. It helps when designing and managing evaluations. It informs evaluation systems. It helps to rethink evaluation methods, design and criteria. It combines methods in evaluation. It addresses delivery risk in uncertain futures. It facilitates collecting and analysing data and reporting issues particularly in challenging contexts. It provides a dashboard that integrates ICT, M&E and managers. It supports emerging technologies including social media, big data, location systems, cyber ethnography etc. It enables evaluation to become foresight. It rewards flexibility and handles complexity. It develops the field of Evaluation to promote resilience and action in critical time. It captures challenges and opportunities for the evaluation field. It addresses some of the dilemmas and trends in professionalism, standards and ethical norms. It advances the theory and applies ethical values to evaluation. It promotes Evaluation Associations as custodians of professional standards. It protects the independence of evaluators and our relevance and responsiveness. It highlights the partnerships and stakeholders who make up the international development aid industry and greatly strengthens the communicating, using and embedding of evaluation.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 116 Evaluating Sexual and Reproductive Health Programs**

**O 339 - Sexual and Reproductive Health Challenges of Adolescent Females in Some Communities of Akwa Ibom, Nigeria**

*J. Inyang*<sup>1</sup>

<sup>1</sup> *Fhi360, Metrics and Research Unit, Uyo, Nigeria*

**Background:** Addressing the sexual and reproductive health (SRH) needs in Adolescent remains a challenge in the rural areas of Akwa Ibom, Nigeria. Hence, this study explores the challenges of adolescent sexual and reproductive health and the perceptions of services to adolescent in some communities of Akwa Ibom State.

**Method:** A qualitative study was conducted among adolescents between the ages of 13–19 years in three facilities of high burden communities (Oron and Uyo LGAs) in Akwa Ibom State to measure approach of adolescence towards services and the perception of Health Care Providers in providing services to adolescence. Data were collected from eight focus group discussion among adolescents who were stratified by marital status and fertility status recruited from the Antenatal care and family planning unit. Six focus group discussion with community health extension workers and 3 key informant interviews with health care providers were conducted. Interviews were tape recorded and transcribed. Data was coded using thematic framework approach and analyzed.

**Result:** Three major sexual and health problems were specifically mentioned during the discussion as challenging to adolescents namely; unintended pregnancies, unsafe abortion and STI/HIV. Healthcare workers often act as a barrier to care by failing to provide young people with supportive, nonjudgmental and youth-appropriate services. There is a need for intervention strategies to tackle issues relating to empowerment, access to resources, improved information and services. Majority of those who abort their pregnancies do so either using local concoctions or visit quacks and unprofessional like the local medicine vendors.

**Conclusion:** This study found that unprotected sex is a common practice among adolescent in Akwa Ibom state. Majority of the participants identified unintended pregnancy, unsafe abortion, and the fear of getting infected with STI/HIV as a major sexual and health concern. Therefore, concerted effort is required by all stakeholders involved in the treatment, care and support for adolescent sexual health in the state to intensify efforts to increase access to treatment and care for young adults in need of reproductive health services and encourage at risk adolescent to get tested for HIV and other sexual transmitted infections.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 116 Evaluating Sexual and Reproductive Health Programs**

**O 340 - Evaluating Digital Tools for Behaviour Change using Big Data**

*C. Lopes*<sup>1</sup>

<sup>1</sup> *Girl Effect, Evidence, London, United Kingdom*

This paper presents an evaluation of the Girl Effect's Springster based on digital data to measure changes on reported behaviour and knowledge related to education and reproductive health of adolescent girls in Philippines. The results of the evaluation shed light on the relevance and impact articles on girls and on the process of change before and after engagement with Springster.

This presentation will discuss the theoretical foundations of the proposed model of impact and questions related to the research design and measurement of digital behaviour. The method consisted of combining spontaneous comments to articles with digital surveys in order to derive indicators of knowledge and behaviour. We will discuss the process of coding the comments, from manual techniques to develop a coding frame to machine learning for classification to code large volumes of textual data.

Although attribution of changes to Springster is not be possible to ascertain due to the correlational design, the question of how knowledge and behaviour of girls change before and after engaging with Springster is answered using structural equation modelling. The results also focus on the segments of girls that are more prone to behavioural change, based on their demographics and previous knowledge. By comparing the results from comment analysis with more rigorous survey techniques, we can unfold questions of validity and reliability of using comments analysis to evaluate impact in a more timely manner.

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 117 Evaluation In Turbulent Times****O 341 - Prospects for Democratic Evaluation in a Polarised and Mediatized Society***A. Hanberger<sup>1</sup>*<sup>1</sup> UCER-Umea university, Applied Educational Science, Umea, Sweden

There is growing worry about increasing polarisation and threats to democracy, but few have discussed how such threats can be dealt with in evaluation. Picciotto (2015) is an exception, having suggested that progressive evaluation can deal with the new evolving context of evaluation. This paper discusses how progressive evaluation and four other democratic evaluation orientations (the elite-democratic, the participatory democratic, the discursive democratic and the market democratic evaluation) can help manage current threats and challenges and how these orientations can be developed. The paper is informed by democracy and mediatization research and democratic evaluation theory.

Obviously, the current conditions for democratic evaluation are radically different from when MacDonald (1976, 1978) first developed the democratic evaluation concept in the 1970s. Today democracy is under pressure, and the status and credibility of scientific knowledge and public media have declined among some citizens, organisations, and politicians. Supposedly, this goes for evaluation as well. Some observers have claimed that the 'open society' and established democratic institutions have weakened (Freedom House, 2018; Keane, 2008). Others have discussed the potential for democratic renewal (e.g. Schmitter, 2015) and point to new forms of political engagement, rising youth participation, social media, and deliberative polls, for example. The paper also pays attention to the mediatization of public policy and governance as this phenomenon has created new challenges for democracy and evaluation.

Whether democracy is described as in decline or in transition affects how we conceive current threats and challenges to democratic evaluation. When rethinking how democratic evaluation can respond to current challenges, the evaluator should consider which narrative of changing democracy to assume. Viewed from a liberal or elite-democratic evaluation perspective (Hanberger, 2006), the notion of decline fits well and the challenge is then to protect current democratic institutions. In contrast, from a participatory or deliberative democratic evaluation perspective, democracy in transition is more suitable. The challenge then expands to one of supporting democratic renewal, for example, through participation and deliberation in new digital settings.

The paper demonstrates that democratic evaluation is poorly prepared to manage current threats to democracy and the mediatization of public policy. Progressive evaluation is the only approach offering new keys to addressing certain current threats and challenges. However, it remains to be seen how well this approach works in practice, whether it can mobilise independent funding, build alliances with advocacy groups, and support democracy as intended. The other orientations have some capacity to manage threats to democracy and support democratic renewal, but need further development. The paper suggests that democratic evaluation could be a constructive tool for maintaining and developing democracy in a polarised and mediatized society if evaluators gain knowledge of threats to democracy, democratic transition, and democratic renewal and, informed by mediatization and democracy research, develop the necessary awareness and competence to deal with these challenges.

The article's findings can be used in deliberating on how to support democracy, addressing the mediatization of policy and governance and what this means for democracy, and developing media strategies.

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 117 Evaluation In Turbulent Times

## O 342 - Trump Our Best Advocate for Evaluation !?

*P. Smits<sup>1</sup>, L. Rey<sup>2</sup>*

<sup>1</sup> U Laval, Management, Canada

<sup>2</sup> Canada

**Rationale:** Since the Patton-Weiss debate in the 1980s regarding the influence of evaluation on decision-making, evaluation has been struggling to reinforce its use. The evaluation community developed approaches sensitive to primary users, participative processes to reinforce mutual learnings about the evaluation and the intervention evaluated, etc. Still a lot remains to be done to mobilize and convince decision makers on the benefits of evaluation. So far, we can not praise ourselves with a stunning success for building an evidence-based society.

This presentation makes the point that we now have a paradoxical great momentum: Mr Trump! We will argue and demonstrate how to draw positive energy out of this turbulent governance style.

#### **Objectives sought:**

- Argue for the unexpected benefits of Trump's style and references to evidence/facts for our evaluation community
- Provide evaluation practitioners and managers with examples of daily Trump inspired-comments/tweets to feed talks around evaluation
- Improve evaluation practitioners and managers skills in demonstrating the importance of key evaluation concepts from Trump-inspired comments/tweets ( especially four concepts: 1. triangulation, 2. credibility of sources, 3. differences between data and judgment, 4. diversity of perspectives from stakeholders)
- Reflect as a community of evaluators upon our own bias regarding turbulent times and opportunities

**Brief narrative and justification:** Turbulent times bring uncomfortable situations that call for a change in our way to analyse and describe new situations, collect data in adjusted manner, and push differently for the use of results.

Turbulent times bring to the fore new leaders, new vocal personalities, unheard voices and eventually new topics, and communication styles. In the person of Mr Trump, a combination of characteristics makes him a turbulent character able to shake the perception of evidence. He employs discourteous vocabulary, communicate with renewed channels via tweeter, openly addresses highly controversial topics, etc.

We introduce the presentation with a provocative illustration of how 'trump-inspired' comments did the job for the evaluation community in putting evaluation in the spotlight. More specifically we present how some politicians launched fact-checking processes, the media counter-argued the sources of evidence used, some evaluation association advocated in the line of post-truth discourse.

Then we present illustrations that this turbulent governance style had repercussions in 4 key areas directly related to evaluation. The 4 key areas are: 1. triangulation, 2. credibility of sources, 3. differences between data and judgment, 4. diversity of perspectives from stakeholders. We draw from media content analysis and tweeter analysis. We selected illustrations for newcomers to evaluation willing to learn about evaluation, and for teachers and evaluation practitioners willing to develop their toolkit and illustrate key evaluative concepts.

We then present how those 4 key areas can be mobilized by the evaluation community to strengthen the visibility of evaluation.

Finally we will open the discussion to the floor and debate upon our personal biases and ethical position, as evaluators, related to the mobilization of unpleasant events for the benefits of evaluation.

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 117 Evaluation In Turbulent Times**

**O 343 - Deliberative Liberalism as a Strategic Approach for Evaluation in Turbulent Times**

*J. Faust*<sup>1</sup>

<sup>1</sup> German Institute for Development Evaluation - DEval, Directorate, Bonn, Germany

Evaluation as a specific form of applied research bridging science and policy-making is faced with serious challenges. Uncertainty about crucial societal developments has been augmenting and at the same time, trust in scientific progress itself is being questioned by an increasing amount of policy-makers and the broader public. Therefore, not only has the scientific task of making policy-recommendations out of past developments become more challenging but the acceptance of applied research for policy improvements itself has been seriously challenged in the past years. Stated in other words, evaluation faces a paradox. On the one hand, in times of mounting ecological stress, violent conflict and social imbalances there is an urgent need for policy improving evaluations. On the other hand, the legitimacy for scientifically based and policy-relevant information has been stagnating at best. Against this background, evaluation not only faces technical challenges about methodological designs or optimization of evaluation processes. More importantly, it faces a normative challenge about how evaluation should be embedded in society to best fulfill its bridging function between scientific inquiry and policy-making.

This paper aims at providing a contribution for overcoming the latter challenge by offering a normative heuristic on how evaluation can contribute for a more enlightened relation between the world of science and the world of policy-making. It does so by applying J. Habermas' seminal work on the relation between science and policy-making to the field of evaluation. In his work, Habermas distinguishes three ideal types of the relation between science and policy-making, namely decisionism, technocracy and (deliberative) pragmatism. In his preferred scenario, Habermas argues that deliberative pragmatism would neither lead to the instrumentalization of science for politics (decisionism) nor to a tyranny of experts (technocracy) but rather to societal progress informed by an enlightened and deliberative exchange of arguments.

Based on this typology, I claim, how processes of public policy evaluation are still constantly being challenged by either decisionism or technocracy and show, how the construction of evaluation processes could benefit from a deliberative perspective. In this regard, evaluation processes should create an inclusive platform for deliberation among different stake-holders along the evaluation process from the early concept of a given evaluation to the dissemination of its results. At the same time, however, a too deliberative perspective on evaluation is not sufficient to guarantee an enlightened influence of policy-making by scientific insights. Instead, deliberative processes in evaluations have to be embedded in a liberal institutional framework. This liberal process on the one hand has to guarantee structural independence regarding stakeholder management, design, conclusions and recommendations of an evaluation, while at the same time respecting the democratic legitimacy of political decision-making.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 118 Rethinking Evaluation

## O 344 - Help! I'm Doing An Impact Evaluation, What Evidence Do I Need?

*J.S. Bayley<sup>1</sup>*

<sup>1</sup> EES, Department of Foreign Affairs and Trade, Yarralumla, Australia

Are quantitative or qualitative methods better for undertaking impact evaluations? What about true experiments? Is contribution analysis the new 'state of the art' in impact evaluation or should I just do a survey and use statistical methods to create comparison groups? This session supports the conference strand 'Rethinking evaluation methods and methodologies' by offering a new perspective on the evidence requirements for undertaking impact evaluations.

Despite their intuitive appeal impact evaluations have yet to gain widespread support amongst politicians, policymakers and administrators. The problem is that they can be expensive to undertake while often yielding equivocal findings with limited policy relevance. Indeed, the quality of evaluations in Australia and New Zealand has been criticised by a wide range of stakeholders including: their respective Auditor Generals; the OECD's Development Assistance Committee; the independent evaluation office in Australia's Department of Foreign Affairs and Trade; and New Zealand's State Services Commission.

So what can be done? Since methodological choices must always be made within a specific context, debates in the professional literature about impact methods can at best only provide partial guidance to evaluation practitioners. The way to break this deadlock is to bypass arguments about methods and instead focus on the evidentiary requirements for assessing causal impacts. Choice of impact evaluation approaches can then be determined by our information needs rather than by our preferred methodology.

This unique session will present a brief summary of the literature on the philosophy and principles of causal analysis, and relate these to some common evaluation models. The presenter will offer three easy to understand evidentiary criteria for reaching rigorous causal conclusions that are grounded in the critical multiplist paradigm.

In this innovative session participants will:

- be exposed to a new way of thinking about impact evaluation
- become familiar with the evidentiary requirements for undertaking impact evaluations
- develop their ability to critique impact evaluation reports
- be provided with resource materials to support their evaluation practices back in the workplace.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 119 Participatory Theory-Based Evaluation****O 346 - Using Social Media for Programme Theory Development: The Role of Emerging Technology in Complex Programme Evaluation***S. Ariss<sup>1</sup>, J. Blackburn<sup>2</sup>*<sup>1</sup> *University of Sheffield, School of Health and Related Research, Sheffield, United Kingdom*<sup>2</sup> *University of Huddersfield, Human and Health Sciences, Huddersfield, United Kingdom*

Crises in areas of society where resilience begins to break down lead to situations in which innovative approaches to improving efficiency are increasingly sought and applied, whilst the resources for engagement in evaluation are put under increasing pressure. It is therefore necessary for evaluators to join in the drive for greater efficiency, by developing more efficient research methods and streamlining the need for engagement with respondents.

Evaluators frequently find themselves operating in environments and dealing with topics with which they have limited prior knowledge or experience. It is therefore critical to engage with experts in the field to collaboratively co-produce the evaluation design and to begin to construct the rationale for activities and expected outcomes. This is often accomplished through the development of comprehensive programme theories or theories of change, for instance in the case of a Realist Evaluation this would be through the construction of Context Mechanism Outcome Conjectures (CMOCs).

Co-producing and validating these programme theories with a wide range of key stakeholders helps to ensure that evaluation findings resonate with people's lived-experiences, which therefore assists with the quality and subsequent uptake of findings and recommendations. However, complex evaluations require input at several stages from many participants for the co-production of programme theories. Traditional methods of engagement, such as discussion groups can be time consuming. This can be a practical barrier to the successful conduct of an evaluation, particularly in areas such as health-service innovation; where organisational resilience is increasingly being tested; resources are increasingly at a premium and there is a key focus on direct service delivery; and members of staff have very little time for anything other than direct service delivery.

This paper describes such an evaluation and how this practical problem of the time-poverty of key stakeholders was addressed through the use of social media. A combination of social media platforms were used to engage with different stakeholder groups, at different levels of specificity. The problem and the development of the solution will be described. Examples will be given of the types of prompts and questions that were asked of different groups, and how these contributed to the development and refinement of the CMOCs to guide the evaluation.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 119 Participatory Theory-Based Evaluation

## O 347 - Key Challenges and Lessons of Making the Best Use of Theories of Change in Thematic Evaluations

*A. Chambel*<sup>1</sup>

<sup>1</sup> UNFPA, Evaluation Office, New York, USA

**Rational and take away:** Overall, participants will learn about the factors that promote effective stakeholder engagement in reconstructing the intervention logic which is at the basis of the evaluation theory of change. The presentation will highlight some of the methodological challenges experienced during the course of the conduct of two concrete evaluation cases which relate to developing complex theories of change; combining different levels of complexity and tackling variations in the case study country contexts.

**Objectives sought:** The abstract focuses on the challenges and lessons of making the best use of Theories of Change in Evaluations. Moreover, the abstract puts special attention on evaluation innovative approaches with a specific emphasis on stakeholder participation informed by human-rights based approach.

**Brief narrative and modality proposed:** Experience to be shared: challenges and opportunities using Theories of Change in corporate thematic evaluations. I will be drawing from two recently completed evaluation exercises: (i) UNFPA support to Census and Data use on Policy Making (2016); and (ii) UNFPA support to prevention, response and elimination of Gender based Violence and Harmful Practices (2018). While the first evaluation looked at UNFPA support to the 2010 census round in 130 countries and its effectiveness in strengthening national capacity for the production, dissemination and use of data to inform decision-making and policy formulation at the country level; the Gender based Violence and Harmful Practices evaluation covered UNFPA support in both development and humanitarian settings as well as the continuum. These two very distinct evaluations offer excellent examples of designing and using theories of change in thematic evaluations.

Topics to be included in the presentation and discussion:

- Reconstruction of the intervention logic and theory of change as a framework for the evaluation design, implementation and reporting.
- Purpose of theories of change: for example, on the GBV and Harmful Practices Evaluation a comprehensive global theory of change was reconstructed to illustrate the dominant approaches to addressing gender-based violence and harmful practices in UNFPA. The purpose was not to test the validity of a 'universal' theory of change, but to map the extent to which mechanisms of change and potential pathways for change are targeted by UNFPA across different contexts, regions and programmes.
- Key challenges facing the development of theories of change, namely: (i) representing the non-linear and complex reality of social change while still representing the key role for UNFPA and taking into consideration the many assumptions at each stage of change; (ii) accommodating differences between how diverse programmes conceptualize how change happens; articulating the intersection between different levels of interventions (individual, family, community, country, cross-border, regional, global); and (iii) distinguishing between different type of (humanitarian and development ) contexts.
- Ex-ante and ex-post theory of change developed at the inception and reporting stages of the evaluation process
- Stakeholder participation, informed by human-rights and gender-equality principles.

Friday, 5 October 2018  
11:15–12:45

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 120 The Relational Aspects of Evaluation****O 348 - Evaluation As A Social Practice***J. Greene*<sup>1</sup><sup>1</sup> *University of Illinois, Professor, Illinois CHAMPAIGN, USA*

**Rationale:** Evaluation is a technical, political, and social craft. The social dimensions of evaluation encompass evaluators' interactions and relationships with multiple, varied stakeholders – from key clients and commissioners of evaluation to program leaders and staff, and to program participants and beneficiaries.

These social and relational dimensions of evaluation have long been acknowledged, albeit quite differently for different evaluation traditions. In large-scale government evaluation, the methodology is typically post-positivist and quantitative, and social relationships are a challenge to be managed or controlled. In case study and field-based evaluation conducted for local or regional organizations around the globe, the methodology is often constructivist and qualitative, or mixed, and social relationships are actively engaged, both as relevant to meaningful data collection and as ethically important. And in evaluations conducted with explicit value or socio-political agendas – including critical, participatory, culturally responsive, and feminist evaluation – social relationships constitute integral strands of the evaluation's purpose and footprint on the ground.

Yet, this social structure of evaluation is rarely included in an evaluation proposal or design. Nor are the relational dimensions of evaluation routinely included in evaluation courses, workshops, or training sessions. So, as a community, we are paying insufficient attention to a substantial – and powerful – strand of our craft. Further, novice evaluators typically enter our profession and "the field" (the sites where the program being evaluated is located) without appropriate understanding of or training in the social and value dimensions of our craft.

**Objectives:** This presentation aims to engage participants in reflection and discussion regarding (a) the importance and power of the social dimensions of evaluation practice and (b) various ways in which we can endeavor to provide a meaningful and practical structure for better understanding the social and relational fabric of our craft.

**Narrative:** The evaluation community spends considerable time and energy engaging with the technical dimensions of our craft – our methodologies and data. And we attend thoughtfully to the politics of the contexts in which we work and thereby of the program we are engaged to evaluate. We also attend to the relationships that are present in an evaluation context, seeking polite and cordial relationships that can ease access to program sites for data collection or facilitate program staff participation in evaluation activities. These are relationships established for primarily instrumental reasons.

But, what about engaging in the social and relational fabric of evaluation on more authentic and caring grounds? A recent book, *Evaluation for a Caring Society*, edited by Merel Visse and Tineke Abma asks this question in a deep and profound way, as it engages the intersections of the professional communities of care and of evaluation. Using this book as one starting point, this paper will wonder aloud about the social dimensions or the social fabric of evaluation practice – how this social fabric gets constituted, how it shows up in evaluation implementation, and how we as evaluators can constitute this fabric in ways that are respectful, inclusive, and caring, while also accomplishing evaluation's substantive and political agendas.

Friday, 5 October 2018  
11:15–12:45

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

## S 120 The Relational Aspects of Evaluation

### O 349 - The Reflective Practitioner – The Road To Innovation

*G.V. Barrington*<sup>1</sup>

<sup>1</sup> *Barrington Research Group- Inc., President, Calgary, Canada*

The domain of Reflective Practice is included in every list of key evaluator competencies yet little attention is paid to what is often deemed a soft skill. In the seminal work by King et al (2001) on competencies, reflecting on practice was ranked 17<sup>th</sup> out of 49 competencies. Why don't evaluators practice reflection?

Becoming an expert has its risks. Schön (1983) suggests that we come to know what to look for and we develop a repertoire of techniques to respond to whatever we find. However, as we become more specialized, we also become less surprised and increasingly automatic in our responses and pay less attention to phenomena that do not fit our pre-conceived categories. As Schön, this can lead to a *parochial narrowness of vision* as well as boredom and burn-out.

At the same time our work is stressful. The environment in which we work is fraught with complexity and our technical expertise is limited by situations of uncertainty, instability, uniqueness, and conflict, a confusing and messy swamp (McDavid & Hawthorn, 2006) where textbook solutions and standards of rigor rarely apply. Seldom do we have adequate resources, time, or control over the phenomenon being studied, yet surprisingly, we are seldom stumped. Kanhehan (2011) suggests that if a satisfactory answer to a hard question is not found quickly, we will find a related but easier question to answer instead. In addition, because we are often tired, stressed, and suffer from the energy depletion common at the end of a project, we do not reflect deeply on the lessons to be learned. Thus, the effectiveness of our evaluation practice is not enhanced significantly from project to project.

Reflective practice is double-loop learning in action. If we think creatively about better outcomes, we can allow ourselves to experiment, innovate, and refocus, expanding our skills and adding value to our work. This presentation will look at how and why we should incorporate reflection more fully into our working lives. What barriers and issues stand in the way of innovation? What questions should we be asking? What reflective strategies can we use, and if we use them, what are the implications? These questions will be addressed as we will look more clearly at the links between reflection, innovation, and action.

Learning objectives include:

- Strengthening our understanding of reflective practice
- Building bridges between reflection and innovation, personally, as evaluators, and within organizations.
- Understanding why innovation is important, what barriers stand in its way, and what strategies can lead to personal resilience, improved practice, and transformative change.

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 121 Pursuing Social Return on Investment**

**O 350 - Social Return on Investment (SROI): Redefining Value and Expanding Possibilities**

*J. Gargani<sup>1</sup>*

<sup>1</sup> Gargani + Company, Evaluation, Berkeley, USA

SROI is a popular – and controversial – method for evaluating the impact that organizations have on society and the environment. It weighs the value of an organization's impacts against the value of the resources required to produce them. Value is, therefore, a foundational concept in SROI, yet it has been criticized by some as underdeveloped. In a paper recently published in *Evaluation and Program Planning*, I proposed a new conceptual model of value. In this presentation, I describe the model and how it can be used to tell evidence-based stories from a variety of stakeholder perspectives. I conclude by acknowledging our growing need to incorporate cost, revenue, profit, and efficiency with other success criteria, and the role that SROI – properly implemented – can play.

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 121 Pursuing Social Return on Investment**

**O 351 - Bringing Evaluation Tools into Impact Investing: Using Theories of Change to Understand the Various Financial Mechanisms Used by Impact Investors**

*N. Verrinder<sup>1</sup>, S. Vaca<sup>2</sup>, A. Wyatt<sup>3</sup>, D. Nixon<sup>4</sup>, K. Zwane<sup>3</sup>*

<sup>1</sup> *Genesis Analytics, Evaluation for Development, London, United Kingdom*

<sup>2</sup> *Independent consultant, Independent, Madrid, Spain*

<sup>3</sup> *Genesis Analytics, Evaluation for Development, Johannesburg, South Africa*

<sup>4</sup> *Genesis Analytics, Evaluation for Development, Oxford, United Kingdom*

Impact investing is becoming one of the largest forces in driving social and environmental change globally. However, how one defines, measures, and communicates this impact is not well defined or consistently implemented. This can prevent investors from making well-informed decisions and allows for "impact washing".

The evaluation community has many tools that could be adapted and used in the world of impact investing. Theories of change allow for better communication of impact, identification of indicators to be measured, and allow for the critical interrogation of logic. The attributes of theories of change could assist in steering the growing force of impact investing towards gathering more investment and achieving greater impact.

This paper investigates the use of theories of change in capturing the logic behind the various financing mechanisms used in impact investing, including payment for success, impact bonds, loans and equity investments. Taking cognisance of the need to show commercial viability, we attempt to marry the logic behind how these mechanisms achieve financial returns with how they aim to drive societal impact. We use innovative visualisation techniques in an attempt to make these easy to understand and useful. While the paper focuses on developing methods of capturing and illustrating theories of change for the various financing mechanisms in the generic, we use real case studies to exemplify how these could be used. The case studies include examples from the housing, microfinance, and agriculture sectors.

Understanding the theories behind how these mechanisms work will provide a starting point to allow for more consistent identification, measurement and communication of indicators of impact.

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 121 Pursuing Social Return on Investment****O 352 - Evaluating Corporate Responsibility in Research and Innovation**A. Wiman<sup>1</sup>, K. Jarmai<sup>1</sup>, A. Martinuzzi<sup>1</sup><sup>1</sup> Vienna University of Economics and Business, Institute for Managing Sustainability, Vienna, Austria

Companies have long been using evaluations to assess the social and environmental impacts of their products and operations. They rely on them to improve their responses to societal demands and reinforce their resilience under changing economic and environmental conditions. In recent years, the debate about what companies can be held responsible for has expanded to include their research and innovation activities. The concept of "responsible innovation" suggests process and output criteria to ensure that innovative products, processes and services serve societal needs and generate no foreseeable negative impact on society or the environment. Definitions of responsible innovation have been proposed mainly in the context of scientific research and remain challenging to be adapted to an industry context. Empirical investigation suggests that even though the concept itself is unfamiliar to industry (Auer and Jarmai 2017), some companies have already implemented practices that operationalize responsible innovation in their research and innovation processes (Lubberink et al. 2017).

The objective of this study is to assess to what extent responsible innovation practices are already being implemented in corporate innovation processes. Previous evaluation approaches (Stahl et al. 2017, van de Poel et al. 2017) have two main drawbacks: First, they approach the issue of responsible innovation from a perspective rooted in the socio-political discourse and are not adequately aligned with the general conditions and processes of innovation in firms. Second, they investigate perceptions, beliefs or levels of awareness about responsible innovation and thus offer little insight into actual corporate practices to ensure responsibility in research and innovation process in firms. To counteract these drawbacks, this study is based on a survey of factual information about observable company practices structured along the research and development phases of corporate innovation processes.

Results will offer empirical insights into responsible innovation practices and their prevalence in companies across industries and sectors. Differentiation between three main phases of an innovation process will allow analysis of corporate practices along the innovation value chain and highlight those areas in which firms have already taken action and those which remain a challenge for companies. Insights will inspire future investigation into evaluation of responsibility in innovation management.

**References:**

- Auer A and Jarmai K (2017) Implementing responsible research and innovation practices in SMEs: Insights into drivers and barriers from the Austrian medical device sector. *Sustainability* 10(1): 1–18
- Lubberink R, Blok V, van Ophem J and Omta O (2017) A Framework for Responsible Innovation in the Business Context: Lessons from Responsible-, Social- and Sustainable Innovation. *Responsible Innovation 3: A European Agenda?*
- Stahl B, Obach M, Yaghmaei E, Ikonen V, Chatfield K and Brem A (2017) The Responsible Research and Innovation (RRI) Maturity Model: Linking Theory and Practice. *Sustainability* 9(6): 1036
- van de Poel I, Asveld L, Flipse S, Klaassen P, Scholten V. and Yaghmaei E (2017) Company Strategies for Responsible Research and Innovation (RRI): A Conceptual Model. *Sustainability* 9(11): 2045

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

## S 121 Pursuing Social Return on Investment

**O 353 - Using Social Return on Investment Analysis to Assist Companies in Optimizing Their Social Value Towards the SDGs***M. Branco<sup>1</sup>, H. Ricardo Lamas Diogo<sup>2</sup>, T. Culhari<sup>3</sup>, F. Serejo<sup>4</sup>, N. Gomes<sup>5</sup>, L. Ribeiro Queiroz de Araújo<sup>6</sup>*<sup>1</sup> *www.mariana-branco.com, EES Board- EES Emerging Evaluators- EES Social Media, Porto, Portugal*<sup>2</sup> *Ramboll, Social Corporate Responsibility-, Salvador da Bahia, Brazil*<sup>3</sup> *Voltália Energia do Brazil, Social Corporate Responsibility, Natal, Brazil*<sup>4</sup> *Voltália Energia do Brazil, Socio-Environmental Analyst, Natal, Brazil*<sup>5</sup> *Voltália Energia do Brazil, Operations Officer, Mossoró, Brazil*<sup>6</sup> *Voltália Energia do Brazil, Social Corporate Responsibility, Socio-Environmental Analyst, Brazil*

**Rational:** The presentation will showcase preliminary results of the SROI analysis of a social project designed by the Brazilian Government, promoted by Voltália (a producer of renewable energy) and financed by the BNDES (Brazilian Bank of Social and Economic Development), who defines a 0,5% social credit of the total approved investment. In Brazil, from 2008 to 2014, the investment by BNDES in renewable energy (wind farms) achieved R\$10.5 billion of reais. The northeast region, where this project is being implemented, has the biggest potential for wind and solar energy projects and the lowest HDI. The project evaluated consists of a community-managed desalination system that provides (1) access to potable water and (2) a fish-farming effluent treatment station to two northeastern villages of Brazil that were affected by a 5 years drought.

**Justification:** Understanding how companies and investors can contribute to the promotion of sustainable development is essential to achieving transformational change through the SDGs and endure social resilience. However, recent studies show that although there's good awareness of the SDGs among companies and investors, there's little consistency in the approach or priorities. Whilst assessing a company's financial performance is straightforward, judging its social impact presents more of a challenge. With the rise of impact investment, a number of methodologies to address these gaps are emerging from different organizations, leading to an unprecedented focus on value for money approaches.

**Objectives:** This presentation will explore one of these approaches: Social Return On Investment (SROI) – a CBA that attributes central importance to stakeholders engagement, materiality and the measurement of social non-market costs and benefits. In particular it will focus on the answer the following questions: 1) how can SROI assist in measuring a company's social impact?; 2) how can it help translating stakeholders views into management practices that lead to social value maximization; 3) how can it contribute to optimize the creation of social value specifically towards the SDGs?

**Preliminary results:** This evaluation is enabling the company to redesign the intervention from a community based perspective; being a framework of principles, SROI incorporates mixed-methods which, in turn, help adjusting the evaluation to real life contexts; the evaluation is allowing the company to better differentiate the outcomes for the community from the outcomes for the firm itself, promoting discussions on transparency and impact maximization. Moreover, as the result of an intense investment in data collection and stakeholders engagement, a range of practical data collection tools for interviews and focus groups have emerged from this study. E.g: an emoji physical likert scale that contributes to more effective outcomes measurement in low income communities; a social impact pizza that when applied in focus groups enables more accurate contribution analysis.

**Questions for the audience:** Optimization for whom? How to deal with complexity? Where do a company's social responsibilities starts and ends? How to manage social value towards the SDG's?

Presentation methods: Icebreaker; handouts such as project presentation, summary of examples and interesting findings; video from the project; short group exercise to trigger the discussion; open materials.

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 122 Evaluation and Theory**

**O 354 - Five, Six or Seven Waves of Evaluation? Across the Western World Since 1965–2018**

*E. Vedung*<sup>1</sup>

<sup>1</sup> Institute for housing and urban research- IBF, Prof emeritus, Uppsala, Sweden

In my 2010 article, I discerned four waves of evaluation that have swept across the Atlantic world since around 1965: the Scientific Wave, the Dialogical Wave, New Public Management Wave and the Evidence-based Wave. All four not only rolled in but also left layers of sediment in public sectors on both sides of the Atlantic.

In the upcoming paper, I will address candidates for a 5<sup>th</sup>, 6<sup>th</sup> and 7<sup>th</sup> wave. They are 5) New Public Governance with cocreation and collaboration across policy and administrative sectors, 6) Ongoing evaluation adopted by the EU during the 2007–2013 program period, and 7) Theory-based Evaluation adopted by the EU during present 2014–2020 program period

**References:**

Vedung, Evert, 2010, "Four Waves of Evaluation Diffusion", *Evaluation: The International Journal of Theory, Research and Practice*, 16(3), 263–277.

Vedung, Evert, 2012, "Four Waves of Evaluation Diffusion", *Evaluation Connections* (European Evaluation Society, EES, Newsletter), January 2012, 4–5.

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

## S 122 Evaluation and Theory

### O 355 - Developing the Field of Evaluation to Promote Resilience and Action in Critical Times: Values in Evaluation: Philosophical, Ethical and Theoretical Considerations

*M. Zounkifrou*<sup>1,2</sup>, *N. Zakariaou*<sup>3,4</sup>, *G. Yaya*<sup>2</sup>, *M. Oussen*<sup>2</sup>

<sup>1</sup> *University of Yaounde II, PhD Candidate/ Economics, Yaounde, Cameroon*

<sup>2</sup> *Health Economics Research Policy and Evaluation for Development Results Group HEREG, Department of planning- management and Evaluations of projects/programs, Yaounde, Cameroon*

<sup>3</sup> *University of Yaounde I, Department of Public Health, Yaounde, Cameroon*

<sup>4</sup> *Chairman of HEREG, Yaounde, Cameroon*

Thinking in value is not thinking about what from the point of view of its evaluation. The concept of "value" is different from "standard" despite similarities in their definition. Thus all social activity is linked to a normativity which is also revealing of its belonging to a class of social activity. Values would therefore be linked to the ethical branch of philosophy, while standards would be attached to the regulatory and legal domain.

Values in evaluation are a set of rules, principles or standards that practitioners must follow for a rigorous and judicious exercise of this activity. Several rules are indeed to be respected; but, in this work, we are interested only in philosophical considerations (objectivist currents, subjectivist and skeptic), ethical and theoretical.

Thus, both theoretical and pragmatic attempts to make objective evaluation are numerous, but in the end they all seem to lead to failure: to reach objectivity seems to be on the order of the chimera of the inaccessible dream. With regard to the subjectivist current, it turns out that the evaluation prepares a decision. This is, of course, not known at the outset; effectiveness, efficiency, fairness and consistency are the different criteria that can be used. The skeptical current, for its part, submits the results of the evaluation to methodical doubts and even of their existence, nevertheless concedes that we constantly emit judgments of values, whatever the function, the purpose of the evaluation.

Moreover, the ethical approach of the evaluation leads us to understand that, insofar as this one interferes with the behaviors of the professionals, it poses a fundamental problem, because there are no preconceived rules that can to be applied and applicable to any context; and any assessment is never neutral in the sense that it affects the people being evaluated, and the results of an evaluation are seen more as a final verdict on self-worth.

Finally, the family of theoretical approaches to evaluation is very broad. Thus, fourth generation evaluation, inclusive evaluation, participatory evaluation, theory-based evaluation, realistic evaluation. The multiplication of theoretical approaches does not guarantee, however, that theory receives all the attention it deserves. This observation is somewhat paradoxical and raises the fundamental question, if any, of the role of theory in evaluation. However, the theoretical value of the evaluation suffers from "benign" negligence on the part of those who practice the evaluation of public policies, programs and projects because the theoretical approaches that are supposed to guide the practice of evaluation do not always seem to be part of it.

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

## S 122 Evaluation and Theory

## O 356 - The Tragedy of Evaluation

D. Briton<sup>1</sup><sup>1</sup> Athabasca University, Centre for Interdisciplinary Studies, Athabasca-, Canada

It's particularly fortuitous, given the nature of the proposed paper, that EES 2018 is being hosted in Thessaloniki, Greece. The paper's aim is to promote resilience and action in critical times, generate new philosophical, ethical, and theoretical positions for consideration, and provide evaluation models for greater societal resilience in response to the critical times of the Anthropocene<sup>1</sup>, perhaps more appropriately the "Capitalocene" (Haraway, 2015). As we progress into the 21<sup>st</sup> century, the possibility of moving beyond the stultifying horizon of the Capitalocene, of "Capitalist Realism" (Fisher, 2015), appears ever more remote. It seems that we have reached "the end of history" (Fukiyama), that "there is no alternative" (Thatcher), and that, as Frederic Jameson (2003) depressingly notes, "it is easier to imagine the end of the world than to imagine the end of capitalism." The paper draws on the work of Greek social and political theorist, Cornelius Castoriadis to trace the emergence of the theoretical perspective that informs Capitalist Realism to the birth of democracy and tragedy in a nascent Athens. For Castoriadis, the birth of Athenian democracy signals the emergence of the first autonomous society, but such autonomy also signalled the end of faith and belief, in favour of universal reason and knowledge. Although both Aeschylus and Sophocles used tragedy to celebrate the contribution of democracy to the development of ethics and politics, it was the latter who "universalized" humanity's new found, democratic autonomy to ALL spheres of human life in an effort to escape the chaos of the Apeiron and the intractable spectre of *hubris*. But the price of escape has proven steep. In fact, what Sophocles proposed has proven to be "a direct recipe for that 'pseudorational pseudomastery' which has gone on to wreak irreparable environmental damage" (Abaffy, 2015). But it wasn't until the European Enlightenment, when Kant totally abandoned Being (ontology) in favour of knowledge (epistemology), that reason reached its apotheosis, firmly establishing humanity the measure of all things. The first irony is that the circumscription of Being that facilitated this transition, far from eradicating *hubris*, delivered humanity firmly into its clutches, and with no seeming escape; the second is that an escape from *hubris* may only be possible through a return to the "flat" ontology of gods, myth, and belief that Sophocles abandoned but Aeschylus promoted. The paper concludes by reconsidering Aeschylus' vision of tragedy by way of a detour through the thought of contemporary speculative realists, Meillassoux ("correlationism"), Laruelle ("anti-philosophy"), and Badiou ("event"). Parallels are drawn between these thinkers and Aeschylus' ontology of the Apeiron in order to generate new philosophical, ethical, and theoretical positions respectful of the creative forces of chaos, cognizant of the ever-present but necessary threat of human *hubris* may emerge to provide truly innovative evaluative models for greater societal resilience.

1. "Humanity's actions have become a new planetary force with accelerating effects on the biosphere. This new era, known as the Anthropocene, calls for new ways of thinking and knowing, and for innovative forms of action" (Gibson et al., 2015).

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

S 123 Evaluation In Extreme Events

## O 357 - Exploiting “Big Data” to Assess Vulnerability and Resilience: Natural Experiments and Observational Research Designs

M. Loevinsohn<sup>1</sup>, N. Mock<sup>2</sup>

<sup>1</sup> Applied Ecology Research, Director, Wageningen, Netherlands

<sup>2</sup> Tulane University, Department of International Health and Development, New Orleans, USA

**Rationale:** Resilience is the capacity of a system, community or society exposed to a hazard to adapt so as maintain an acceptable level of functioning and structure. A good deal of research and evaluation assesses that capacity hypothetically, in relation to characteristics assumed to confer resilience. This presentation will discuss and illustrate research designs that can assess actual performance in the face of hazards by exploiting existing large data sources.

Individual or multiple data sources may contain information relevant to the exposure of people, individually or in aggregate, to a hazard; outcomes of or responses to exposure and contextual factors. Hypotheses about relationships between these can be tested statistically using cross-sectional designs. These are appropriate where the hazard is an idiosyncratic (individually experienced) shock or stress. A weakness is that the direction of causality may be ambiguous.

Where the shock is covariant e.g. an extreme event affecting large numbers and areas, a natural experiment design is possible. The event can be framed as an uncontrolled experimental intervention which defines periods before and after exposure and people or groups that are exposed and unexposed or differentially exposed. Stronger causal inferences about exposure-outcome relationships can be drawn than with cross-sectional designs.

**Analysing the Malawi Famine 2001–03 as an uncontrolled intervention:** Three data sources were used to assess vulnerability and resilience:

- Nation-wide survey of rural households' food assistance needs at the Famine's peak, by district
- HIV infection from surveillance surveys just before and after the Famine, by woman and rural/non-rural site
- Nation-wide household survey which documented migration and crop cultivation

Multilevel analysis revealed that HIV prevalence at rural sites increased sharply in the districts hardest hit by hunger, consistent with people's accounts and ethnographic studies of women being pushed further into survival sex. At the non-rural sites (towns and cities), which had been spared much of the hunger, HIV prevalence **declined** in an inverse relationship with the extent of rural hunger in the surrounding district. Several strands of evidence indicate that this resulted from rural women, predominantly young women dependent on agriculture, migrating in search of food and work. Because prevalence in villages was roughly half that in towns and cities, rural women moving there diluted prevalence. More women moving from districts where hunger was greater created the inverse relationship.

The analysis revealed a systemic failure to address vulnerability to food price volatility, leading to destructive coping responses – survival sex and distress migration. Their pervasiveness and impact on HIV's dynamics had not been previously appreciated.

Resilience was also revealed. Districts where more households had taken up cassava, which tolerates drought better than maize, had lower maize prices – a collective benefit – less hunger and less change in HIV.

This natural experiment sheds light on complex, multi-scale processes which an RCT would be unable to address. The increasing frequency and severity of different extreme events suggests an expanding potential for natural experiments. Widening access to quality big data sources and enhancing skills in their analysis will be key to realizing that potential.

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

S 123 Evaluation In Extreme Events

## O 358 - School-based Data Collection in the Context of an Earthquake: Evaluation of a Comprehensive Sexuality Education Program in Mexico City

S. Makleff<sup>1</sup>, J. Garduno<sup>2</sup>, I. Zavala<sup>2</sup>, P. Rivera<sup>2</sup>, F. Barindelli<sup>3</sup>, C. Marston<sup>1</sup>

<sup>1</sup> London School of Hygiene of Tropical Medicine, Public Health and Policy, London, United Kingdom

<sup>2</sup> Fundación Mexicana para la Planeación Familiar- A.C. Mexfam, Evaluacion, Mexico City, Mexico

<sup>3</sup> International Planned Parenthood Federation/Western Hemisphere Regional, Gender Sexuality and Rights, New York City, USA

**Rationale:** When evaluating social interventions, data collection is often constrained by the institutions and settings in which it is conducted, particularly in relation to environmental disasters, shifting political contexts, and other unpredictable factors. These commonly lead to midstream changes in data collection strategies, which are rarely examined beyond the limitations section of evaluation reports or publications. It is important for evaluators to share data collection challenges and potential strategies to address them, and to reflect on how midstream adjustments to data collection strategies may influence the rigor and utilization of evaluation findings.

**Objectives:**

1. Describe the challenges of school-based data collection in Mexico City as affected by the September 2017 earthquakes.
2. Report on the adjustments to data collection timelines and strategies due to earthquake-related school closures.
3. Examine how these adjustments may influence the perception of rigor and potential utilization of findings.

**Narrative and justification:** International Planned Parenthood Federation/Western Hemisphere Region (IPPF/WHR), Fundación Mexicana para la Planeación Familiar (Mexfam), and the London School of Hygiene and Tropical Medicine (LSHTM) collaborated on a study evaluating the school-based comprehensive sexuality education (CSE) intervention developed by Mexfam and implemented by the organization's "Gente Joven" ("Young People") program. The mixed method process evaluation examines whether and how CSE contributes to the prevention of dating violence, shifts in gender norms, and more equitable relationships among 14- to 17-year-old public school students in Mexico City.

The evaluation uses a longitudinal quasi-experimental design in one intervention and one control group. Data collection methods include a pre-post survey; longitudinal qualitative interviews with intervention participants; endline focus groups in both intervention and control groups; endline in-depth interviews with intervention participants; and endline focus groups with teachers and "Gente Joven" health educators.

In 2017, the CSE intervention was piloted from February to June and implemented from September to December. On September 7<sup>th</sup> and 19<sup>th</sup> there were two powerful earthquakes in Central Mexico, interrupting the CSE intervention and baseline data collection for the evaluation. Several weeks of school closures led to uncertainty about whether the semester would continue and if the intervention and evaluation study could be completed.

When the school reopened, Mexfam completed baseline survey implementation, although the timing differed significantly across study participants, and there were additional challenges such as youth leaders influencing levels of participation of other students. Mexfam continued the CSE intervention with some adjustments. First, to address anxiety among students after the earthquake, the "Gente Joven" educators repeated the process of building trust in the group and setting ground rules. The contents of the intervention were also compressed into fewer

hours, while keeping all critical contents. These implementation changes were compounded by challenges recruiting participants, particularly young men, and scheduling interviews. In this presentation, I will reflect on how the adjustments to intervention implementation and the data collection timeline in the emerging post-earthquake context may have influenced evaluation rigor and the potential utilization of findings, in this case increasing the relative importance of the qualitative data.

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

## S 123 Evaluation In Extreme Events

**O 359 - Methodological Decisions for Evaluating Rural Community Resilience Investments in Malawi: Maximising Rigour and Utilization**K. Wilson<sup>1</sup><sup>1</sup> LTS International, Consulting, Edinburgh, United Kingdom

**Rationale:** Agriculture employs 86% of the national workforce and contributes 38% of Malawi's GDP and 90% of its export earnings. However, much of the rural economy is highly vulnerable to both climate and market shocks and changes. During the period from 1970 to 2008, Malawi experienced more than 40 weather-related disasters, with ten major flood or drought events, affecting an average of 2.3 million people each and causing annual crop losses of US\$149 million on average. Estimates of the future impacts of climate change in Malawi are also large, with additional costs equivalent to an additional 2% of GDP per year by 2030–2050, and rising thereafter. Many development initiatives aimed at building resilience have been short-term and focused on small numbers of interventions.

The Enhancing Community Resilience Programme funded by the UK's Department for International Development, the Royal Norwegian Embassy and Irish Aid took a unique approach offering combinations of up to 11 interventions at household and community levels, supported by national level policy advocacy. The programme also invested in a 5-year independent monitoring, evaluation and learning function which delivered a suite of mixed methods evaluation activities, including participatory learning activities, cost-benefit analysis, process tracing and quasi-experimental impact assessment.

In deciding on methods to evaluate resilience programmes, commissioners and evaluators must select methods which are practical, affordable and which will be used by stakeholders to make crucial design or implementation decisions.

**Objectives:** This paper will explore the decision-making process for identifying appropriate methods to respond to the particular measurement challenges posed by this type of resilience programme. It will explore the cost, practicality and quality considerations that informed these decisions. In particular, the paper will focus on the application and utility of three different methods. 1) Participatory learning journeys for grassroots workers; 2) quasi-experimental impact studies and 3) Intra-household assessments focused on gender equality.

**Narrative:** This presentation will explore an evaluation designed to improve the design and implementation of interventions promoting the resilience of rural households and communities in Malawi. It will describe a range of mixed-methods evaluation techniques which promoted both rigour and utilisation and which feedback to improved practice in an ongoing resilience intervention in a highly vulnerable context. A focus on gender equality is a key feature of all investments made by the UK's Department for International Development and the approach to incorporating this in the evaluation design can also be explored. The partnership between a UK and Malawian evaluation firm is a key feature of this programme and can be highlighted in the presentation.

NB: It may also be able to link this presentation in a round-table to a related presentation made by Dr Barbara Befani on the application of *Process Tracing with Bayesian Updating* to a policy influencing activity funded under the same *Enhancing Community Resilience Programme*. The attendance of a DFID staff member involved in the evaluation commissioning is also feasible and could add an additional perspective to a potential round table.

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 091 Humanitarian Interventions and Evaluation

## O 360 - Listening to What Children Need, Think and Recommend

*O. Kinda*<sup>1</sup>

<sup>1</sup> Save the Children, Monitoring Evaluation Accountability and Learning MEAL, Dakar, Senegal

Since the ratification of the Convention on the Rights of the Child there has been a slow shift within international development agencies towards acknowledging the rights and capacities of young people to participate in decisions that affect their lives. Current needs assessment processes, including the UN Multi-Cluster Initial Rapid Assessment (MIRA), do not ensure the systematic consultation and participation of children (Harcourt, Deborah Sargeant, Jonathon Sargeant, 2012). This commonly results in the outputs of such needs assessments failing to include the unique perspectives and experiences of children. As such, how do we effectively support children and give them opportunities to make their voices heard and influence projects and programmes outcomes?

Evaluators do not usually have the time and the space required to fully explore critical issues related to child and youth participation. Providing such opportunity and platforms for discussions and reflections is essential to taking things forward and providing stakeholders including children and youth with knowledge, new perspectives and ideas.

Recognising this gap, child-centered agencies such as Save the Children have developed for instance a child-specific needs assessment methodology that enables girls and boys to voice their needs and opinions, using methods appropriate to their ages and evolving capabilities. Other initiatives incorporate key recommendations and standards for Children's Participation.

This paper analyzes participation of children as a means to enhance evaluation and research. It explores tools and methodologies that have been tested and validated in recent experiences of Save the children with its implementing partners including children in West and Central Africa. It comprises data collection tools and accompanying guidelines for: assessment planning, data collection, data analysis, report writing and dissemination of results. Type of participation, benefits of participation for children, ethical issues, ways of working with children are explored. Knowing that continuous efforts are needed to effectively involve children in the world of evaluation, the study reviews good practices and strategies, and comes up with propositions of ways to address identified challenges and to strengthen participation.

Evidence-focused literature review and recent studies of Save the Children including participation mapping, evaluation and research reports, progress reports and country annual reports will inform the paper. Finally, and crucially, this presentation enables agencies, evaluation professionals to better support children to realise their rights enshrined in the United Nations Convention on the Rights of the Child (UNCRC).

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 125 Mixing Methods in Evaluation**

**O 361 - Evaluation and Statistics: The New Partnership**

*S. Beaujean<sup>1</sup>, N. Gomes<sup>2</sup>, C. Omes<sup>3</sup>*

<sup>1</sup> ARTEMIS Information Management SA, Department for International Cooperation, Luxembourg, Luxembourg

<sup>2</sup> Ministry of Foreign and European Affairs, Directorate for Development Cooperation, Luxembourg, Luxembourg

<sup>3</sup> LuxDev, Evaluation and Knowledge Management, Luxembourg, Luxembourg

Public policies with solid and well-defined indicators monitored from the outset are still to be largely widespread. However, evaluable policies based on sound statistics strengthen governance and as a consequence support the development of more resilient societies. Programmes or projects can miss clear baselines or target values. In addition, the application of the Paris Declaration principles on development effectiveness – in particular alignment – in programmes and projects can sometimes weaken the embedment of result framework.

By default, programme evaluations may be based mainly on qualitative analysis because of the lack of results frameworks including SMART (Specific – Measurable – Achievable – Relevant – Time-bound) indicators. The combination of both quantitative and qualitative analysis is not always a common practise.

Combining quantitative and qualitative approaches strengthens the objectivity of the analysis, enlarges the analytical scope and increases the evaluation's credibility. Having an evaluation based on reliable statistics enhances the acceptance of findings and recommendations by stakeholders.

The challenge is how to make sure that statistics are reliable? There is a need to thoroughly analyse the methodological approach employed to process the indicators – statisticians refer to this as metadata – and to ensure that international standards and methods defined for the compilation of statistics are applied. Knowledge exchange and collaboration between evaluators and statisticians is crucial to overcome reliability problems and use the real power of statistics.

Although statistics may not always be gender-specific, the use of data disaggregated by sex is crucial. The analysis of disaggregated data enables the identification of phenomena specifically related to women/girls or to men/boys thus allowing for gender-specific recommendations.

It is important to raise awareness within the evaluation community that evaluators have a role in advocating for better evaluability of policies, programmes and projects. Evaluators and statisticians need to exchange experiences on how to identify reliable data sources, on how to properly interpret the data and on the analytical/statistical skills to integrate in a multidisciplinary evaluation team.

The objective of this paper is to raise awareness about the added value and challenges of combining quantitative and qualitative approaches in evaluations and about the importance of collaborating with statisticians to understand data and to know how to interpret it.

The Luxembourg Ministry of Foreign and European Affairs, LuxDev and Artemis propose a comparative analysis based on two practical evaluation cases implemented in 2017–2018. They are based on the OECD/DAC evaluation criteria and on official statistics or statistics assessed as "reliable" by the evaluation team in collaboration with statisticians. The two cases are:

- A retrospective evaluative study of 15 years of cooperation between Luxembourg and El Salvador where no detailed results framework existed;
- A peer-review process of a regional HIV/AIDS programme implemented by Civil Society Organisations where a results framework with both quantitative and qualitative indicators existed.

This comparative analysis illustrates:

- (i) the added value of combining quantitative and qualitative approaches in evaluations;
- (ii) potential ways to overcome challenges with regards to results frameworks;
- (iii) the difficulties met to collect and interpret data and how to overcome these.

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 125 Mixing Methods in Evaluation****O 362 - Integrating Qualitative Research Methods into Economic Evaluation***S. Gobin<sup>1</sup>, S. Foley<sup>1</sup>*<sup>1</sup> Palladium/MEASURE Evaluation, Data- Informatics and Analytical Solutions, Chapel Hill- NC, USA

Economic evaluation methods are often thought of as solely quantitative in nature. Use of qualitative data to support partial economic evaluation methods, like costs analysis, offers a novel opportunity for enriching cost analysis outputs. Methodologically, what does it look like to implement a mixed-methods approach to economic evaluation during data collection and analysis in practice? This paper hopes to answer that question by reviewing the approach used to design, collect, and analyze data from six USAID-funded programs implementing Orphan and Vulnerable Children case management across six countries in Sub-Saharan Africa. Researchers from the USAID-funded MEASURE Evaluation Project used a parallel data gathering approach, conducting in-depth qualitative interviews with more than 150 program staff and case workers across national, sub-national, and community levels. Retrospective program financial and beneficiary data were collected as deemed appropriate for the same implementation levels. Content analysis of researcher interview notes was used for the qualitative data, while a step-down costing approach was used to categorize, assign, and allocate costs. Qualitative data were drawn from when making decisions around categorizing, assigning, and allocating costs. Program beneficiary data was then used to calculate cost per beneficiary and qualitative results were used to explore and interpret quantitative data findings. Researchers found that conducting a mix-methods economic evaluation was feasible in the context of international development program evaluation and served as an innovative methodological approach that resulted in richer overall findings, allowed for the exploration of additional research questions linked to cost, and provided a new lens for interpreting cost per beneficiary data.

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 125 Mixing Methods in Evaluation**

**O 363 - Visualizing Change: Mixed Method Data Analysis  
in Evaluation with ImpactMapper**

*A. Pittman*<sup>1</sup>

<sup>1</sup> *ImpactMapper, Founder, Arlington, USA*

Evaluators, foundations and nonprofits possess rich data on social change strategies that are working and those that are not. A majority of this data exists in qualitative form – in the form of grantee reports, stories of change, reports, interviews, field mission notes, etc. Too often this data remains unanalyzed because there are not easy to use tools that allow for custom tracking of outcomes, along with other quantitative and grantmaking data. Moreover, often we privilege quantitative metrics in the sector, but they don't tell the whole story, especially when thinking about longer term societal change processes, like increasing human rights and eradicating racial and gender inequalities, which takes longer than 2–3 years to achieve. This means as a sector we may be missing out on key social change trends and strategies that are working.

This session showcases diverse ways that ImpactMapper, an online tool to analyze and visualize mixed method data, has been used for monitoring, evaluation and research purposes. Three case studies will be shared:

1. M&E of a foundation's grantee portfolio (Oak Foundation's, Issues Affecting Women Program),
2. Meta-analysis (UNDP, Annual Development Report Assessment) and
3. Perception Surveys (PAWHR network, Collective Grantmaking Assessment).

This session will help support evaluators' ability to more easily analyze the data they have in a variety of formats, from word, excel to creating custom surveys. The tool supports evaluators in a number of analytic activities, all while supporting the creation of beautiful data visualizations and dashboards to enrich your reporting.

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 126 Countering Socio-Economic Hardships

## O 364 - Cultural Development and the Overcome of Poverty: Evaluation of the 'Servicio País Cultura' Program in Valparaíso, Chile

A. Mallo<sup>1</sup>

<sup>1</sup> Ministry of Culture and Innovacional consultant in Argentina and Chile, Development of Arts and Culture-International Policy Evaluation in Latin America and the Caribbean, Buenos Aires and Santiago, Argentina

The Servicio País Cultura Program was implemented between 2011 and 2015 by the National Foundation for the Overcoming of Poverty (NFOP), together with the Ministry of Culture of Chile (MCC) and local municipalities, as part of a sociocultural policy that addresses from a multidimensional approach the challenge of overcoming poverty, giving a crucial role to the cultural development of the participating communities.

The Program had a presence throughout the whole national territory and to the date of this research it had not reported any programmatic evaluations that aim to understand the processes of its implementation and the perception of the subjects involved, both within the institutions as well as in the communities.

Thus, the research seeks to assess the implementation of the program, according to the perceptions of the community participants, the members of the management team at the NFOP, MCC and the Municipality, taking as a study case the Municipality of Puchuncaví, located in the region of Valparaiso in Chile.

Using a qualitative research design, I assess the logical model and the Program's theory of change in the particular context of this territory, giving account of its strengths and weaknesses, from the perspective of the main actors involved.

As a result, while the program shows high internal consistency and it is accepted as a real contribution in general areas, it also remains insufficient in the resolution of the main conflicts identified before implementation. In particular, two points of conflict were found: on the one hand, it lacks consideration of the territorial conditions in the intervention plan, and on the other, there are substantial problems in the institutional coordination among the NFOP, MCC and the Municipality of Puchuncaví.

Conclusions are based on the importance of including territorial pertinence and organizational coordination in the processes of design and implementation of the Program. Also, a relevant issue to be taken into account in further program management is to encourage an 'evaluation culture' within the institutions that hold part in the Program, in order to reinforce the presence of evaluation criteria in the policy cycle as a whole. Lastly, recommendations regarding these elements are drawn, as to improve the management of this national sociocultural policy.

Therefore, this evaluation intends to achieve a better understanding of the cultural dimension in the social phenomenon of poverty and, hence, improve public policy strategies aimed to address it. Indeed, the assessment provides a set of lessons that can be used to the proper functioning of the program, as well as to enhance the relations between communities, institutions and territories.

This experience shows how evaluation can contribute to large development goals, such as the overcome of poverty and the improvement of general life conditions, but also how the assessment process itself strengthens community bonds and their linkages with local governments and agencies, which together make societies stronger and more resilient in the long run.

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 126 Countering Socio-Economic Hardships

## O 366 - Evaluation of the Growth and Poverty Reduction Strategy Paper III of Cabo Verde

*E. Sarmento*<sup>1</sup>

<sup>1</sup> Nova Business School of Management and Economics, Novafrica, Lisboa, Portugal

The **Growth and Poverty Reduction Strategy Paper (GPRSP) III** is the strategic document encompassing Cabo Verde's transformation agenda, which describes how the country aimed at fulfilling a competitive and sustainable economy and continue towards the path of poverty reduction during the years 2012–2016. During this period, the country has considerably hit by the financial crisis, and still remains vulnerable to economic and environmental shocks.

The **evaluation objectives** focus on the macro and microeconomic performance and achievements for the five-strategic axis (Governance, Human capital, Infrastructure, Private Sector, Global Nation), the seven clusters (Agro-business, Information Communication Technology, Financial services, Maritime Economy, Creative Economy, Aero-business, Agribusiness, Tourism), the range of programs and reforms implemented during the period, the national planning system and the monitoring and evaluation system. The evaluation also assessed the resource mobilisation through the annual budgets from 2012 to 2016, carried out with the support of the PEFA methodology.

The **methodology** is one of mixed methods (qualitative and quantitative). It uses Theory-based Evaluation and Contribution Analysis with participatory and gender-based principles. The evaluation criteria follow the OECD-DAC's criteria; relevance, effectiveness, efficiency, impact and sustainability. Three other criteria were brought into the assessment framework to complement the analysis: "Partnerships and leverage", "Managing for results" and "Cross-cutting issues".

The gender assessment of Cape Verde's rural communities was done through fieldwork and focus groups with women-led families, case studies on gender and sociological analysis of the role of women in society, the workforce and trends of equality and economic empowerment over time. Furthermore, two surveys were elaborated: (i) on training, employment and employability and (ii) the national systems of planning.

**Findings:** The current challenge for Cape Verde is economic and fiscal consolidation to contain macroeconomic vulnerabilities. The economic and financial viability of state owned companies was fragile and should be reviewed in light of the country's financial situation.

Relative to employment, there have been no substantial improvements. Unemployment continues to affect mainly young people aged between 15–34, with a level of education between primary and secondary education, and those living in urban areas. The precariousness of employment remains high. The proportion of resources earmarked for employment promotion and the amount of Programs set over the implementation period are not consistent with the importance given to employment promotion in the DECRP III strategy.

Substantial inequalities persist in Cape Verdean society and especially among the poor, and this should not be neglected in the design of public policies.

Findings also point to gender inequality persisting in Cabo Verde. Women suffer from high levels of gender with substantial violence and discrimination against women in general. The position of women remains socially and financially lower than that of men. The areas of higher employability and salary continue to be occupied by men. Women's rights have not improved, despite the passage of important legislation such as gender-based violence law, and the Law that permits abortion in public hospitals.

On institutional sustainability, it is necessary to reinforce the culture of accountability.

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 127 Inclusive Economic Development**

**O 367 - Sustainable Development through Alternative Energy:  
The Implication of Socio-Economic Development Programmes**

*L. Ndala*<sup>1</sup>

<sup>1</sup> *Johannesburg, South Africa*

South Africa has recently implemented the world's most successful renewable-energy independent power producer procurement (REIPPPP) in the world. In an attempt to streamline the economic development, the REIPPPP programme requires independent power producers (IPPs) implement specific guidelines that speak to job creation, enterprise development and socio-economic development. In line with relevant legislation, IPPs should assess and implement socio-economic development programmes within a 50 kilometre (30 miles) radius of a project site as part of their power producer agreement.

This qualitative research study explores the outcomes and possible impact of implementing socio-economic development programmes in complex, rural and remote environments in the Northern Cape, South Africa. Moreover, the study looks at how the evaluation findings has influenced strategy development and implementation. Over 100 participants participated in focus groups, surveys and semi-structured interviews to determine effectiveness, sustainability efficiency and impact.

The findings demonstrate that socio-economic development programmes in complex, rural and remote environment often require a longer time frame to measure outcomes and impact. In addition, programmes need to be agile and receptive towards the evolving needs of the community through an asset-based community development process.

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

## S 127 Inclusive Economic Development

### O 368 - Analyzing the Program Theories of the Policies Against Poverty and Social Exclusion. The “Income of Inclusion” in Italy

*C. Torrigiani*<sup>1</sup>

<sup>1</sup> University of Genoa, Di.S.For., Genova, Italy

Between 2008 and 2016, in Italy, the population at risk of poverty or social exclusion increased from 25.5 to 30 percent and people in material deprivation ranged from 7.5 to 12.1, with a peak of 14.5 in 2012 (Eurostat). These evidences, together with the historical absence in Italy of a structural measure against poverty, they explain the attempts made in this direction by the governments Renzi and Gentiloni, as requested among other things by the commitments to the achievement of the Europe 2020 objectives. Renzi Government introduced in 2015 the so-called “Sustain for Active Inclusion” (law 208/2015). Based on that attempt and on a “delegated law containing rules concerning the fight against poverty, the reorganization of services and the system of social interventions and services” (law 33/2017), Gentiloni Government introduced in 2017 the “Income of Inclusion” (legislative decree 147/2017).

The introduction in Italy of a measure to fight poverty, already experimented in the period 1998–2000 in the form of a “minimum income for insertion” (law 237/1998) clashes with a welfare model still based, on one hand, on the occupational status rather than on citizenship (Esping-Andersen, 1990) and, on the other, on the centrality assigned to the family as a social safety net (Naldini, 2002; 2003) with a residual role of the public actor. A welfare model characterized by a strong functional and distributive distortion (Ferrera, 2012) compared to other European countries. The centrality of this issue has been highlighted by the recent electoral success of the 5 Star Movement, which has largely based its election campaign on the promise of a generous “Citizenship Income”, not without confusion among voters about the real meaning – and the feasibility in Italy – of such a measure. The electoral success of the Movement has in fact concerned, particularly, the southern regions, characterized by higher rates of poverty and unemployment.

The purpose of this contribution is to present a punctual evaluative analysis of the “Income of Inclusion”, highlighting the actors involved in its implementation and their roles, the organization envisaged for the implementation of the measure, between vertical monetary and information flows and the need for horizontal coordination between services and operators at the territorial level, compared with actual implementation. The analysis will put in evidence the dimensions of considerable complexity deriving from the territorial articulation of the Italian welfare in a strongly “fragmented” system (Kazepov and Barberis, 2013), in which the hypothesis of uniform implementation of a measure such as the “Income of Inclusion” clashes with the reality of highly inhomogeneous regional governance systems.

To this aim, on one hand, will be adopted analytic keys useful for the evaluability assessment (Wholey, 1979), such as organizational plan and service utilization plan as a specific articulation of the program process theory (Rossi, Freeman & Lipsey, 2004). On the other, reconstructing the program impact theory, will be problematized the hypothesis of change on program beneficiaries (Weiss, 1998) keeping in mind the variety of “mechanism, context, outcome” configurations (Pawson & Tilley, 1997) as a guiding criterion.

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 127 Inclusive Economic Development

## O 369 - Evaluating Market System Development Programmes: The Value of Sector Knowledge

*C. Smit*<sup>1</sup>

<sup>1</sup> Genesis Analytics, Evaluation for Development, Johannesburg, South Africa

Evaluations of programmes that aim to influence a market system are complex – they require deep contextual understanding, an analysis of changes at the micro, meso and macro/eco-system levels and often span multiple geographies. Certain evaluation approaches are also preferred – most commonly theory-based evaluation and contribution analysis. While all evaluations require a combination of evaluation and sectoral expertise, evaluations of market system development programmes need to bring together and find a balance between understanding relevant evaluation approaches, the market system development approach (which is built on fundamental principles of market facilitation and the achievement of scale and sustainability) and the sector/system in which the programme is operating. Different degrees of emphasis are placed on each of these three areas by different clients.

This presentation hopes to explore the challenges associated with striking the right balance between all three.

Genesis Analytics has conducted a number of evaluations of market system development programmes, a number of which will be showcased during the session:

- i) a five-year programme aiming to extend digital financial solutions to 1 million smallholder farmers in Kenya, Tanzania and Zambia,
- ii) a review of a grant to a non-profit organisation with the goal of greater investment in affordable housing and housing finance throughout Africa, and
- iii) an intervention aimed at creating employment opportunities for unemployed young people in Mombasa, Kenya.

This session will draw on those experiences to discuss sound approaches to these types of evaluations, what can go wrong and what can be done to structure evaluation teams and workplans to avoid these pitfalls.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 005 Institutionalising Evaluation

## O 370 - The Role of Parliamentary Evaluation Forum for Creating Resilient Society

*A. Prasad Pokharel*<sup>1</sup>

<sup>1</sup> *Legislative Parliament of Nepal, Nepal Government, Kathmandu, Nepal*

Parliamentarians are major stakeholders of evaluation as their role is crucial in making and shaping the policies of public concerns including but not limiting to the ones related to resilience societies. To support the policy related issues of evaluation and encourage parliamentarians to make evidence-based policies, National Parliamentary Forum on Development Evaluation Policy in Nepal (NPFDEPN) was established in 2014 which was the first National Forum of Parliamentarian on Evaluation globally. The NPFDEPN as an evaluation association has been playing important role in promoting evaluation nationally and internationally in partnership with other stakeholders. Internationally, it played crucial role in hosting the Evaluation Week in 2015 in the parliament of Nepal where several far-reaching agenda were launched including EvalAgenda2020, formation Global Parliamentary Evaluation Forum, and EvalNetworks such as EVALSDG, EvalGender+, EvalYouth, and EvalIndigenous. While partnering with National VOPE, Community of Evaluators and government of Nepal, NPFDEPN developed Integrated National Evaluation Plan 2020 and also contributed significantly in preparing the Equity-Focused and Gender-Responsive Monitoring and Evaluation Bill of Nepal that has emphasized on paying special attention to the issues of vulnerable people including their resilience to natural and human created disasters. The role of the Forum has also been significant in making the humanitarian policies to support the affected population by the 2015 earthquake that shook Nepal brutally while killing almost 9000, injuring over 22000 and displacing internally about 95,000 people. In the parliament of Nepal, there is a Good Governance and Monitoring Committee to monitor and review parliamentary affairs and development interventions undertaken by parliamentary bodies in order to promote transparency and accountability of parliamentarians toward their constituents. The NPFDEPN while partnering with the Committee organised workshops for the members of the Committee to enhance their understanding on monitoring and reviewing interventions and make evidence-based policy.

These are a few examples what a parliamentary forum for evaluation can contribute to creating enabling environment in evaluation field building. The Evaluation Forum can contribute to making the work done with efficiency and effectiveness as tendency of the bureaucrats, especially in the developing world, is to listen more attentively to the issues brought to them by the parliamentarians than by others. As the parliamentarians are a link between the people and government and they are answerable to the questions and concerns of the people including the vulnerable people, who are also the voters of the parliamentarians, their understanding and engagement in evaluation help them formulate right policies and ask development planners prepare customized plans and programmes specific to the needs and demands of different kind of the people, not leaving the vulnerable behind. However, things are not as simple as said. The situation is more complicated under the turbulence, armed conflict, crisis and disaster. The proponent of this session will share the experience about the challenges faced and mitigation strategy adapted and encourage participants to provide feedback what evaluation approaches would be even better than what has been presented.

Friday, 5 October 2018  
11:15–12:45

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 128 Promoting Use of Evidence in Government**

**O 371 - Evaluation and New Public Management: Substitution or Complementarity?**

*O. Dolder*<sup>1</sup>

<sup>1</sup> *INTERFACE Policy studies and University of Lucerne, Department of Political Science, Lucerne, Switzerland*

Various reforms of New Public Management have been implemented in Switzerland in the last 20 years. At the same time, there has been continuous development of policy evaluation. The proposed paper clarifies the relationship between evaluation practices and NPM. Analysing public administration units of some Swiss cantons in the health and education domain, the paper examines the mutual influence between two core instruments of NPM, namely, performance and impact targets with performance and impact indicators, and lump sum budgeting, and two important elements of evaluation practices, namely, evaluation activity, and institutionalization of evaluation. This examination is relevant from a theoretical and a practical point of view: There has been lively theoretical debate on whether NPM reforms reinforce policy evaluation or NPM performance indicators that focus on administrative outputs substitute for (replace) policy outcome evaluation.

Friday, 5 October 2018  
11:15–12:45

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 128 Promoting Use of Evidence in Government****O 372 - Institutionalizing the Evaluation as an Improvement Mechanism for Public Interventions***M. Aioanei<sup>1</sup>*<sup>1</sup> *National School of Political Science and Public Administration, International Relations and European Integration, Bucharest, Romania*

The paper addresses the topic of evaluation utilization in the context of the European Cohesion Policy. The aim of my research was to identify the factors that influence the level in which, through evaluations, the interventions financed within the Romanian System of Structural Instruments (SSI) can be better orientated and more efficient. First, using the theory of Patton, I have determined the level of **instrumental utilization of evaluation** within the Romanian System of Structural Instruments (SSI) in the financial period 2007–2013 (+2), which included 7 Managing Authorities, each for one Operational Program (OP) coordinated at the national level. Through this approach the research aims to create a comprehensive view upon the utilization of evaluation phenomenon identified in the Romanian public administration. Second, using data collected from experts involved in the management, implementation and evaluation of the OPs, I have developed a comparative analysis which emphasizes the importance of several influence factors identified in the process of enhancing the role of evaluation in the public sector, with stress on **conceptual utilization**. Therefore, the analysis is focused on the mechanisms that have positively influenced the decision-making process and the stakeholders, in order to create a wider acceptance of the need of evaluation in the public sector and also to enhance the role of evaluation's conclusions and recommendations within the management of the OPs. Third, the paper includes also a comparative overview of the previous financial cycle and the current one in terms of changes done within the evaluation public system.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

## S 140 New Developments in Monitoring for Evaluation

### O 373 - Sustaining Monitoring and Evaluation in Health Care

*M. Teisen*<sup>1</sup>

<sup>1</sup> *Capital Region of Copenhagen, Topmanagement strategies, Frederiksberg c, Denmark*

Any management idea or recipe is challenged, when it's no longer part of the latest fashion in management design. Even a highly accepted and well-functioning management tool can be challenged by the fact, that it was invented too long ago, by another management team, or in a slightly different arena. The consequence can be lack of proper implementation, misinterpretation of results and lack of performance.

**Problems for large Organizations:** For large organizations like hospitals, the lack of sustainability of management ideas can be a large problem. The capital Region of Denmark employ 40.000 highly skilled people. One single hospital can employ as much as 12.000 doctors, nurses' radiologist etc. In such an organization it takes years to implement a coherent system such as Monitoring and Evaluation. Before data-flows, follow up mechanisms and learning loops are in place and well accepted it could easily take six to eight years. That's a long time to wait.

**The aim of the paper:** This paper will present ideas for sustaining the concept of monitoring and Evaluation in such organizations. The paper is based upon experiences from The Capital Region of Denmark, and its hospitals, serving more than two million people in an around Copenhagen. The aim is to establish a framework of options for further sustaining monitoring and Evaluation concepts, drawing from lessons learned through the close work with management and employees in the large organization. Also, the paper draws upon lessons learned during the performance of the Voluntary Peer Review (VEPR) of the management officer.

**Monitoring and Evaluation in The Capital Region of Denmark:** The Monitoring and Evaluation system in The Capital Region of Denmark is based upon indicators decided at the regional (top-management) level management as well as indicators decided on all other management levels.

Management and employees are meeting regularly to discuss performance on selected indicators. Further work for improvement of the performance is discussed. Monitoring and Evaluation contributes both to accountability and to organizational learning.

**Working on sustaining monitoring and Evaluation:** Some of the tools used at sustaining monitoring and evaluation in the Capital Region of Denmark are mentioned below.

Being a politically governed entity, the Capital Region of Copenhagen is facing both the opportunities and challenges connected to political decision-making. However the experience is, that politicians are largely appreciating the detailed performance information, and that the political involvement by itself contributes to sustaining the system by creating a substantial demand for data.

Focusing both on improvement and on accountability at the same time can be a challenge. However, by developing a coherent improvement model, and by establishing an education for working on improvements based on performance information, the challenges are met and the capacity for working on the system is enhanced.

Finally, going through the Voluntary Evaluator Peer review (VEPR) has provided a great pool of knowledge and a comprehensive action plan for further work on the system has been worked out. The paper will draw on many of the conclusions from this plan.

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 129 Evaluation Standards and Credentialing

## O 374 - Evaluation Standards Development and Use: The case of the International Development Evaluation Association (IDEAS) and UNICEF

D. Schroeter<sup>1</sup>, M. Tarsilla<sup>2</sup>

<sup>1</sup> Western Michigan University, School of Public Affairs and Administration, Kalamazoo, USA

<sup>2</sup> UNICEF, West and Central Africa Regional Office, Yoff Dakar, Senegal

**NOTE: Hanna, as we discussed in our email exchange on March 29, 2018, this paper belongs into the larger panel on “Evaluation Standards and Competencies for Resilient Societies” (EE18–0298)**

The purpose of this panel is to engage presenters and the audience in a critical discussion on two key topics that appear relevant to the resilience of the evaluation profession in the future, both within and outside of Europe. First, what regional and global evaluation standards and competencies have been developed so far and how? Second, what type of relationship exists between the existing evaluation standards and the evaluation competency frameworks currently available?

The panelists, who have over 80 years of combined evaluation experience, will discuss their direct experience in developing and using evaluation standards and competencies in numerous countries and organizations. Regardless of their distinct perspectives on these issues, panelists will share their experience and expertise in relation to the following questions:

- How have they contributed to the development of evaluation standard?
- How have the developed standards been adapted over time to the needs of the different evaluation communities which they were supposed to serve, both regionally and globally?
- How are evaluation standards used, and by whom?
- What is the relationship between evaluation standards and evaluation competency frameworks?
- How have evaluation standards been updated to reflect the recent developments in theory and practice (including emerging barriers and opportunities)?
- Which actors are the most engaged in standards and competency development?
- What is the role of VOPEs, universities, and other stakeholders in the development of evaluation standards?
- To what extent and how are standards and competencies independent from, and relevant and responsive to the needs of their users?

The session chair will aggregate key comparative documentation from the participating (and other) organizations to set the stage for facilitating debate and discussion between the panelists and audience.

The panel is equally relevant to novice and more seasoned evaluation practitioners, managers, commissioners and users as it raises their awareness about the opportunities and intricacies of developing evaluation standards and competencies worldwide. Furthermore, the panel aims to clarify (through the use of case studies) the extent to which evaluation standards and competency frameworks are furthering the utility, feasibility, propriety, and accuracy of evaluation endeavors in a number of organizational and cultural settings. In doing so, the panel is intended to advance the current discourse on how to tap into the opportunities and address the challenges associated with the innovation of the existing evaluation standards. Lastly, the panel is

expected to provide the audience with a number of concrete scenarios in which evaluation professionals “resilience” was tested and amply accrued in the past. In an attempt to enhance the audience capacity to deal with the currently tumultuous political climate characterized by “fake news” and “alternative facts,” the panel will demonstrate how critical discussions across cultures and openness to “otherness” not only foster adaptiveness in evaluation practice but also promote more effective action towards increasingly resilient and equitable societies.

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 129 Evaluation Standards and Credentialing****O 375 - Evolution of the Canadian Evaluation Society's Professional Designation, 2010–2018: Implications for Resiliency of the Field***N. Kishchuk<sup>1</sup>, B. Gauthier<sup>2</sup>, G.V. Barrington<sup>3</sup>, H. Cummings<sup>4</sup>*<sup>1</sup> *Program Evaluation and Beyond Inc., Montréal Québec, Canada*<sup>2</sup> *Réseau Circum Inc., Gatineau, Canada*<sup>3</sup> *Barrington Research Group, Calgary, Canada*<sup>4</sup> *Harry Cummings and Associates Inc., Guelph, Canada*

In 2010, the Canadian Evaluation Society (CES) introduced a voluntary professional credential, based on demonstration of a combination of education, experience, and professional competencies and with an ongoing professional learning requirement. Currently, about 20% of CES's membership have been awarded the Credentialed Evaluator designation, and another 12% are in the process of applying; survey data suggest another 25% are considering applying. This presentation will be made by individuals responsible for the Credentialing Program's implementation and management during this period: Natalie Kishchuk is currently CES Vice-President and responsible for the program; Benoît Gauthier is CES Past-President and formerly Transitional Vice-President responsible for the program; Gail Vallance Barrington was formerly CES Vice-President, managed an evaluation of the CE designation and chairs a review of the CE competencies; and Harry Cummings is currently CES President. The presentation will first briefly describe the Credentialing Program's its history and operations. It will then outline uptake patterns among Canadian and international CES members during 2010–2018 as well as among employers and commissioning organizations, and their implications for future evolution. The presentation will also describe the Credentialing Program's performance against targets and provide a viewpoint on how it has been contributing to the strength and resiliency of the evaluation profession in Canada. Finally, it will comment on some unexpected and emerging trends and concerns. The presentation is related to Strand 3 "Developing the field of Evaluation to promote resilience and action in critical times," especially "Dilemmas and trends in professionalism, standards and ethical norms."

Friday, 5 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 129 Evaluation Standards and Credentialing

## O 376 - Evaluation as an Agent of Change

*L. Back*<sup>1</sup>

<sup>1</sup> Lucien Back, Independent consultant, Amsterdam, Netherlands

**Rationale:** In its quest to promote accountability and learning, evaluation theory and practice has developed codified approaches and methods to ensure credibility, usefulness and timeliness of its findings and analysis. In the process, evaluations have become increasingly uniform and predictable limiting their aspiration to be innovative and challenging and acting as agents of change with a view to make societies more resilient.

**Objectives sought:** This contribution to the EES Conference 2018 will explore institutional constraints to independence and innovativeness of evaluation functions, e.g. organizational embeddedness, pressures on evaluation managers and commercial conditions to be met by external evaluators. It will explore a realistic way forward for institutions, evaluation managers and evaluators to make evaluations more performing in terms of useful contributions to resilient societies.

**Narrative / justification:** Under institutional constraints described above, terms of reference are often negotiated compromises that need to address concerns of multiple stakeholders in organizations resulting in a too high level of ambition and too many questions to be answered. The problem is often compounded by a lack of realism in terms of timelines and budgets.

Faced with these constraints, the community of evaluators has increasingly resorted to the common denominator shared by public and private sources of funding: the quest for results, with a bias to confirming the status quo. Logical framework approaches and their more recent manifestation, theories of change, press reality into causal chains and relegate factors and dimensions that do not fit into external conditions and hypotheses.

These approaches are useful to demonstrate to what extent intended results have been achieved in terms of outputs, outcomes and impact and may also yield evidence if these results have been produced at a reasonable cost. They are less performing to detect unintended consequences or to answer the question, whether the activity or programme under review made sense in the first place. They also leave too little space to views and aspirations of stakeholders at the bottom of the power paradigm usually conditioning the evaluation function.

In an ideal world, the independence of evaluation functions in organizations should be strengthened. This will only be possible, if authorities commissioning and funding evaluations do not only use them as self-serving processes justifying the status quo, but open up to be challenged in more or less fundamental ways. However, the resulting shift in the power balance is unlikely to happen in a major way.

Under these circumstances, evaluation managers and evaluators should at least make attempts to break out of the pre-established mould and make sure that the following questions are addressed as honestly as possible: Did the activity or programme made sense in a broader context and also from the perspective of voiceless stakeholders? To what extent were there or will there be unintended consequences? Were there or can there be alternative ways of doing things?

Emphasizing such questions would strengthen the learning dimension of evaluation (and its users) and make evaluations less predictable, more innovative and challenging and more useful to strengthen the resilience of societies.

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 130 Performance Monitoring and Rating Systems****O 377 - Constructing the Outcome Monitoring System for Portfolio of Biotechnology Programs with Mixed Methods and Benefits Diffusion Model**T.Y. Liu<sup>1</sup>, C.C. Chang<sup>1</sup>, Y.C. Lin<sup>1</sup>, A.Y. Lo<sup>1</sup><sup>1</sup> Science & Technology Policy Research and Information Center, National Applied Research Laboratories, Taipei, Taiwan

With the increasing scarcity of STI program budget and rising accountability pressure, Taiwan's government are pressured to assure the better allocation of STI investment. Due to the high uncertainty of STI programs, the STI funding agencies are increasingly aware of the managerial needs to improve the programs management before the long-term benefits unfold. Therefore, Taiwan's funding agencies need a program portfolio evaluation tool to inform decision making in the short-term. Our proposed composite rating system is adapted from ATP's Composite Performance Rating System (CPRS) designed to evaluate the distribution of whole portfolio for completed projects. The most pressing issue regarding the available reported quantitative indicators is their absence of underlying rationales, timeframe, and context. Therefore, we utilized the logic models to redefine the available quantitative data, delineating the benefits diffusion channel through direct users and indirect users, constructing the primary benefits dimensions, and selecting the proxy indicators. The Portfolio of Biotechnology Programs were divided into commercialization, extension, and R&D programs, which respectively contained different measurement dimensions combination. Our proposed portfolio performance rating system consists of six dimensions: "innovation capacity building", "knowledge creation", "knowledge dissemination", "technology commercialization", "industrial diffusion", and "informing policy making". The selected quantitative indicators (through case study) are translated into numerical total score by means of algorithm based on qualitative expert interviews, and are translated into a four-star rating system. Through this methodology, we can monitor the ongoing programs to inform the managerial interventions by means of annual output data to signify the outcome progress of the monitored programs. The performance distribution of 86 annual data from completed programs are as follows: 8%, 4 stars; 20%, 3 stars; 20%, 2 stars; and 52%, 1 or 0 stars. The data as 4 stars and the data as 3 stars contained 75% showed greater scores in the commercialization dimension. The performance of biotechnology programs was distributed in line with recognition of field experts. The results of the rating system provide an easy-to-grasp tool to communicate portfolio performance and progress of program goals. For the further step, we need to separate immediate outcomes from intermediate outcomes in order to monitor and improve the ongoing programs. Overall, the portfolio performance offered by the rating system could inform the government and funding agency of ongoing managerial intervention and coordination by focusing on the users/partners in the specific benefits diffusion trajectory.

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 130 Performance Monitoring and Rating Systems****O 378 - Getting it Right: A Mix of Evaluation Methodologies for Complex Research for Development Programs**M. Guertin<sup>1</sup>, T. Schuetz<sup>2</sup>, S.K. Gaffney<sup>3</sup><sup>1</sup> CIMMYT, Research and Partnerships, Texcoco, Mexico<sup>2</sup> CGIAR, Research Program on Climate Change- Agriculture and Food Security, Munich, Germany<sup>3</sup> CIMMYT, Monitoring, Evaluation and Learning Specialist, Mexico city, Mexico

Monitoring and evaluation in a multi-layered organization setup has become a challenge in terms of complexity. Often faced with a wide range of requirements by different stakeholders, the monitoring and evaluation systems need to be aligned and integrated with various strategic plans and broader strategies. The demand for systematic collection of performance information and its management adds to challenge in making the monitoring and evaluation system efficient. This context requires a whole new, creative thinking of evaluation methodologies.

This paper presents the steps taken and lessons learned in the establishment of a mix of innovative and more robust evaluation methodologies to address accountability, informed-decision making, adaptive management and learning needs by many governance and organizational entities. It shows two case studies of a research for development center and program under the CGIAR, a global research partnership for a food-secure future, i.e. International Maize and Wheat Improvement Center (CIMMYT), and Research Program on Climate Change, Agriculture and Food Security (CCAFS).

CGIAR, as the system level, developed a new Strategy and Results Framework (2016–2030), which sets out ambitious targets aligned with the Sustainable Development Goals (SDGs). To reach these targets, a portfolio of a number of research programs (e.g. CCAFS) was put together, managed in partnerships across the 15 centers (e.g. CIMMYT). In absence of a blueprint on how to build a robust monitoring and evaluation system, and based on existing assessment of the strengths and weaknesses of previous systems developed for large organizations, CCAFS and CIMMYT approached the challenge by focusing on the results-based management principles as the foundation for their monitoring, evaluation and learning (MEL) systems.

The innovative approach taken by CCAFS and CIMMYT was choosing different elements from existing MEL systems, adapting them, putting them conceptually together into an efficient and robust MEL. CCAFS and CIMMYT selected the following components: performance framework; harmonization of impact pathways and theories of change; indicators and baselines; reflective spaces and activities; reporting; assessment and bonus allocation; institutional transformation; management information system. The additional element of cross-cutting dimensions, such as gender, youth, capacity development, were applied across all the elements. This allowed for a more comprehensive and holistic set of elements to meet the principles of RBM and to be able to support the programs in meeting their contributions to the SDGs.

We recognize the following systems, which we borrowed inspiration and components for the center and program MEL systems: OECD DAC countries, Asian Development Bank, United Nations Development Program, World Bank Results-based Management Frameworks; International Aid Transparency Initiative standards; Sustainable Development Solutions Network; OECD.

Both at center and program level, the MEL systems were introduced in a very decentralized and participatory manner, including a wide group of experts, including upper and middle management, who had to take responsibility and play a role in the establishment and implementation for the related processes. The standardization of processes was supported by the investment into an online tool that guides people intuitively through the processes along the results-based logical causal chain.

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 130 Performance Monitoring and Rating Systems

## O 379 - The Use of Social Media in Monitoring and Evaluation: Opportunities and Challenges

*M. Canares*<sup>1</sup>

<sup>1</sup> *Step Up Consulting Services, Evaluation and Strategy, Tagbilaran City, Philippines*

The growth of mobile users in developing countries has been unprecedented and this occurred only in the last five years. In most of these contexts, access to the internet is done via mobile phones and largely by using social media sites, even creating the perception that social media sites are outside the internet. While there is a significant amount of literature in monitoring and evaluating social media initiatives, there is little less attention on how social media can be used to conduct monitoring and evaluation. This paper seeks to fill this gap by proposing a conceptual framework in operationalising the use of social media in conducting monitoring and evaluation of development projects.

Through actual case studies in Southeast Asia that illustrate how social media can be used in monitoring and evaluating development projects, this paper argues that there are inherent advantages in the use of social media to gauge the outputs and outcomes of development projects, and in a few cases, impact. This includes efficiency of data collection, the ability to mix different qualitative data to provide evidence for project achievements (e.g. photographs, poll, videos, and qualitative rankings), and the feasibility of large-scale sample data collection with less costs. However, it also posits several disadvantages – lack of verification, biased self-reporting, the potential for results to be skewed towards people with technology access, among others.

Nevertheless, the paper argues that there are enabling conditions that can facilitate the use of social media to capture project results. At the same time, sufficient controls can be installed to ensure that the risks of inappropriate recording and reporting can be abated. While still at its nascent stage, the paper argues that with the right condition and conditionalities, there is potential for social media to be used as an effective tool in the conduct of monitoring and evaluation. The paper proposes an operationalisation framework to achieve this end.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 131 Big Data and Emerging Technologies in Real World Context**

**O 380 - Benefits of Using Big Data in Evaluations and How to Overcome Technical Constraints**

*R. De Luis Iglesias<sup>1</sup>*

<sup>1</sup> Ipsos MORI, Policy and Evaluation Unit, London, United Kingdom

The huge potential for using innovative methods relying on big data to feed into research studies and evaluations is increasingly being recognised. The growth of digital data is providing access to information in difficult to reach locations, and generates real time information on issues such as the cost of food, availability of jobs, gender equality, access to health care, or quality of education, to list but a few examples that could be relevant for an evaluation. It also has the potential to overcome some constraints posed by traditional methods such as surveys and interviews. For example, the analysis of social media activity can incorporate data from all over the world, thus avoiding cost limitations or low participation rates.

However, the use of 'big data' by evaluators is still very limited. This is partly due to the fact that data scientists are not normally evaluators, and most evaluators are not yet fully aware of the potential of emerging technologies, or do not know how to access, extract and process data for inclusion in their evaluations. The objective of this session is to showcase a few examples in which the author and her colleagues at Ipsos have used big data in evaluations or research studies. The examples will illustrate the potential of innovative methods in evaluation and how technical barriers to obtain data can be overcome. The methods explored in these evaluations include social intelligence, bibliometric analysis, text analytics, and passive measurement.

For instance, we will show how we used social intelligence to measure the outcomes of a monitoring report on education across the world. Analysis of social media through the platform Synthesio helped the evaluation team to explore the volume of conversation and its location, both in terms of channels and geographic location, who the influencers were and the sentiment of those commenting on the report. In this evaluation, bibliometrics and text analytics were also used to provide information on the elements of the report that were most valuable to academics and policy-makers.

Other examples include a project for the Food Standards Agency in which we analysed consumer interactions with Food Business Operators over Twitter with the tool Method52, developed by the University of Sussex, and the evaluation of an academy programme in the British health-care system to improve quality of care and patient outcomes, in which Ipsos used text analytics with an IBM Modeller software to assess a series of impact stories collected by the programme on individuals who had been through the Academy's programmes.

Given the great potential of big data, evaluators should explore new tools to exploit these data and combine and triangulate findings with traditional tools. Innovative tools can provide access to information that otherwise would be unavailable, or available only at a high cost, and help triangulate findings from surveys and interviews that might be subject to bias. However, the use of big data also poses some ethical issues that must be acknowledged and considered in the evaluation design and the data storage and treatment.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 131 Big Data and Emerging Technologies in Real World Context

## O 381 - The Best of Both Worlds? Integrating Traditional Evaluation and Data Science Methods

*K. Bertermann*<sup>1</sup>

<sup>1</sup> *Girl Effect, Evidence, London, United Kingdom*

Digital platforms for social impact offer exciting opportunities for data analysis and understanding users. However, it can be confusing to make sense of 'live' and significant amounts of digital data. Using Springster, Girl Effect's global site for adolescent girls, this paper briefly describes a number of traditional and data science tools which are available for analysing digital data, including Google Analytics, social media analytics, online surveys, and comment analysis. This paper will offer an analytical framework for making sense of digital data, placing each data source in relation to its usefulness for understanding users' engagement with a site, participation on the site and impact of the site on a user.

**Objectives:** The objectives for the paper are:

- 'De-mystify' digital data collection methods
- Propose ways in which emerging data science methods can complement traditional data collection and analysis methods

**Relevance:** NGOs and social impact initiatives are increasingly using digital platforms to inspire civic participation and influence knowledge, attitude and behaviour change. In contrast to traditional interventions, digital platforms provide continual live data streams in a variety of formats, including back end data from Google Analytics, online surveys, social media data, and user generated data, including comments and stories.

These data sources offer a wealth of data, but the potential of the data to contribute to a compelling evaluative framework is often overlooked. Using Girl Effect's Springster mobile site (aimed at 14–16 year old adolescent girls) as a case study, this paper describes a framework which can be used to 'make sense' of digital data in order to add depth and nuance to monitoring and evaluations. The framework capitalizes on the immediacy of digital data and thus includes "engagement" and "participation" metrics as well as traditional impact metrics. Engagement metrics describe users' activity on the site participation metrics describe users' active interactions with other users and the site itself. Impact metrics include standard measures of knowledge, attitude, and behaviour change.

In addition to a measurement framework, digital data also offers the opportunity work with large-scale data sets. In order to analyse data at scale, data science methods such as social network analysis, machine learning and predictive analytics are useful. However, while data science is crucial for large-scale data sets, more 'traditional' data analysis methods are still crucial, particularly for understanding changes in users' knowledge, attitude, and behaviour.

The paper will also describe how digital data collection and analysis offers opportunities for disaggregating data for specific user groups. For example, some highly engaged users might be mobile phone owners, rather than borrowers. Understanding these user behaviours helps contextualize and ground the data.

While digital data holds great promise and opportunity for providing rapid monitoring and evaluative results, digital analysis still benefits from "analogue" or offline methods as well. While relying mainly on digital analysis methods, the Springster measurement and analysis framework also includes periodic face to face sessions with some users, in order to sense-check and refine the analytical framework.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 131 Big Data and Emerging Technologies in Real World Context****O 382 - Redefining Monitoring and Evaluation Systems for Development effectiveness of Watershed Development Programs***R. Bangalore Krishnaiah<sup>1</sup>*<sup>1</sup> ANTRIX Corporation limited, Department of Space, Bangalore, India

Monitoring and Evaluation (M&E) in Developmental projects has assumed greater importance and has become an essential and indispensable tool for tracking the progress and performance of the project, streamlining the interventions and for assessing the impacts. The traditional monitoring systems are not as robust as desired, since it focuses on basic physical and financial criteria and excludes data on implementation quality and impacts. With the advancement in technology, modern tools like Remote sensing, (RS), Geographical Information System (GIS), Global Positioning System (GPS), Management Information system (MIS), play a vital role in different facets of monitoring and evaluation of developmental projects. These ICT tools has advantages over traditional approach, for gathering and managing information on project inputs and outputs, tracking project impacts and outcomes to answer questions about progress against broad developmental indicators. The availability of high resolution, multi-sensor and multi-frequency satellites with the capability to provide unbiased data has facilitated the generation of inputs required for taking up developmental works as well as a means for monitoring the implementation and assess the impacts.

Integrated Watershed development program is considered as an effective tool for addressing problems of drylands and is recognized as potential engine for agriculture growth and alleviate poverty in India. With financial assistance from World Bank, "Sujala" was conceived as sustainable poverty alleviation cum watershed development program and was implemented in a massive scale encompassing 1270 villages covering 518,000 ha and 350,000 families in Karnataka, India for a period of seven years. This program utilized the cutting edge technologies to plan, prioritize interventions, monitor progress and assess impacts.

The uniqueness of the Monitoring & Evaluation (M&E) approach evolved and adopted in this project lies in the fusion of modern technologies like Remote Sensing (RS), Geographical Information System (GIS), Global Positioning System (GPS) and Management Information System (MIS) with conventional monitoring system, which made a robust M&E system to provide the state-of-the-art for tracking the project progress, provide evidence based outcomes and impacts. The high quality data and reports generated through technological tools helped the project to identify bottlenecks early on, undertake mid-course corrections and shift the project direction many times. These modern tools helped in assessing the impacts following the difference in difference estimates between pre and post and control data to depict positive changes like increase in crop yield, cropping intensity, biomass, ground water level, house hold income, milk yield, etc, reduction in runoff, migration, school dropouts etc. at the end of the project. Significant impacts in all the three dimensions of the sustainable development viz economic, social and environmental aspects were demonstrated due to biophysical and social interventions coupled with adoption of technology for web enabled governance and to enhance transparency/accountability in the project.

World Bank has acknowledged this M&E approach adopted and operationalized in Sujala project as "Model of Excellence" and conferred "Global Best Practice" Award. The project has been recognized with many International and National Awards and is being replicated.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 132 Real Applications of Realist Evaluations****O 383 - A Realist Evaluation Approach to Studying Impacts of Dance Based Interventions on Parkinson's Patients**S. Sridharan<sup>1</sup>, A. Nakaima<sup>2</sup><sup>1</sup> University of Toronto and St. Michaels Hospital, The Evaluation Centre for Complex Health Interventions, Toronto- Ontario, Canada<sup>2</sup> St. Michaels Hospital, The Evaluation Centre for Complex Health Interventions, Toronto, Canada

Realist evaluation seeks to answer questions around 'what works for whom and under what conditions.' Despite this promise, there are very few examples in the literature of realist evaluations that have rigorously reported outcome pattern data. This expert lecture will discuss how theory based evaluation such as realist evaluation can help study the impacts of dance based interventions. This lecture discusses a realist impact evaluation designed for Dancing with Parkinson's (DWP), a community organization that has been delivering dance classes for individuals with Parkinson's disease for 9 years. Although Parkinson's disease (PD) is a neurodegenerative disease characterized as a movement disorder, ironically dancing seems to not only help people with PD to initiate and coordinate their movement while dancing but also helps them to better manage living with PD. Realist evaluation is theory driven and the realist mantra is "what works for whom and under what circumstances." The first phase of the evaluation explored mechanisms with participants (what is working for them) and linked the mechanisms to self reported outcomes. A second phase of evaluation is planned employing a quasi-experimental design further exploring impacts with a comparison group of non-dancers. Also we plan to follow DWP participants over time using a method called Qualitative Longitudinal Research (QLR).

Research methods for testing pharmaceuticals/drugs are well established and the health system relies heavily on pharmaceutical treatments based on the proliferation of evidence of drug effectiveness generated from research. However evidence of non-drug treatments are less available and often viewed as less robust (also there are fewer research dollars spent on non-drug treatments). As a field evaluation has the opportunity to test methods, approaches and study designs that can help build the evidence supporting non drug treatments such as community interventions like dancing for many reasons including that community interventions often are cheaper and have fewer negative side effects than drug interventions. A realist evaluation approach can have wide applicability to study arts based interventions. We think that with this focus on configurations of context, mechanisms and outcomes, our experiences would have relevance to a number of other situations

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 132 Real Applications of Realist Evaluations

## O 384 - In Search for a Realist “Fiat”: Lessons from Operating Realist Evaluation with Limited Budgets

A. Devaux<sup>1</sup>, T. Delahais<sup>1</sup>, A. Williams<sup>1</sup>

<sup>1</sup> Quadrant Conseil, Research, Paris, France

During her keynote address at the last EES congress, Oxfam GB's Claire Hutchings regretted that research in evaluation mostly discussed sophisticated evaluation designs, the equivalent of fine-tuned “Ferraris”, whereas in real life evaluators need “Fiats”: practical, cheap and easy to handle methodologies.

The realist approach is one of these Ferraris (Pawson & Tilley, 1997). Despite a straightforward inquiry agenda summarized by a search in evaluation for “what works, with whom, and under what circumstances”, the realist evaluation approach and its core concepts such as mechanisms, are subject to regular “tuning” and definition updates (Astbury & Leeuw 2010; Pawson 2013)

This approach, except in public health, is often perceived both by evaluation commissioners and providers as mainly adapted for high budget and elaborate evaluation or research designs, mostly used after other methods have already been used to assess the impact of a programme (Robert & Ridde, 2013; Salter & Kothari, 2014). Hence, examples of use of this approach in Europe and particularly in France remain scarce.

Yet, does the realist approach need to always be so sophisticated? Can it be adapted to small evaluation missions with limited budgets and yet still correspond to its core concepts?

As researchers and practitioners curious to test this approach, we seized the opportunity when relevant to test the realist approach in such contexts. Our team adopted and adapted the realist approach in several evaluative contexts, including two full-fledged evaluations in France in 2017. It was challenging using the realist approach with these evaluations, as they were operated with rather low budgets, within a short period of time and with clients with no prior experience in evaluation practice. These evaluations provided two very distinct contexts of implementation. The first one assessed a school initiative to promote the recycling of light bulbs amongst primary school pupils and their family and the second one assessed the impact of financial support from a private foundation to biomedical researchers.

These two experiences enabled us to strengthen a “Highway code” on why, how, and under which circumstances we could use realist evaluation, what kind of results we could expect from this approach and how the demonstration of causality it provided was convincing to commissioners.

First, we formalized ways to assess which evaluation demands were potentially suited for realist evaluation even with limited budgets (1). Second, in order to address both questions about identifying the impact of a programme and understanding why they occurred, we combined the realist approach with a theory-based evaluation within a common evaluation grid and restricted the use of realist approach to selected causal packages of the theory of change (2). At last, we built knowledge on how to present results about mechanisms in a clear and synthetic way for the client and translated them in practical recommendations (3). In light of these experiments, we identified obstacles and established situations in which the Fiat realist evaluation would not be feasible (4).

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 132 Real Applications of Realist Evaluations

## O 385 - Combining Realist with Developmental and Utilization-Focused Evaluation Methodology: Conflict or Symbiosis?

*S. Ariss<sup>1</sup>, J. Blackburn<sup>2</sup>*

<sup>1</sup> University of Sheffield, School of Health and Related Research, Sheffield, United Kingdom

<sup>2</sup> University of Huddersfield, Human and Health Sciences, Huddersfield, United Kingdom

The central theme of this paper is the combination of popular evaluation approaches that have incorporated complexity concepts. It will reflect on over ten years of applying and teaching complex evaluation methods in a wide range of settings. This work has largely been inspired by an early experience of producing a final report for a two-year programme evaluation, which was largely treated as inconvenient and comprehensively ignored. An experience that I am sure is shared by many evaluators. During this evaluation, some approaches proposed by Michael Quinn Patton in his (1986) book 'Utilization-Focused Evaluation' (UFE) were adopted, but this was 'too little, too late' to make a noticeable impact on the usefulness of this piece of work.

In the following decade, an intensive programme of complex evaluations was carried out, in which UFE approaches were adopted alongside Realist Evaluation methods to increase the focus on 'usefulness'. The development of these combined methodological approaches, and their application in complex settings, was accelerated in 2011 and 2013 by the publication of Patton's 'Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use', and Pawson's 'The Science of Evaluation: a Realist Manifesto'.

There are numerous differences between the approaches of Pawson and Patton; perhaps most notably the former's emphasis on scientific methodology and the latter's focus on pragmatics and usefulness. Key similarities are the central importance of; developing theories of action, and the conceptualisation of causal linkages in complex settings.

The combination of Realist Evaluation approaches with other approaches has been a contentious issue, which has often divided the community. This is particularly the case when epistemological or ontological foundations appear to be incompatible. For instance, the ongoing debate concerning the combination of Realist methods with trials-based methodology has been a rich source of controversy.

This paper suggests that, far from being incompatible, there are greatly beneficial synergies to be gained from combining both these approaches. Many years' of experience in applying and teaching Realist Evaluation methods has highlighted issues around prioritization, boundary-setting and decision-making that can potentially be resolved with this proposed combination of approaches. Perhaps the greatest advantage is the potential for a systematic focus on usefulness to underlie the theoretical development process of CMOCs (Context, Mechanism, Outcome Conjectures). Thereby, helping to define the role of the evaluator and providing focus and boundaries for CMOC development; preventing 'theory-creep' and evaluator bias.

The successful combination of these methodologies can help to ensure that evaluations are scientifically rigorous and focused on answering questions that are of greatest importance and relevance for key stakeholders, thereby supporting the resilience of societies.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 132 Real Applications of Realist Evaluations****O 386 - Although Realistic Evaluation has Improved Ascertaining Causalities, it is Still Fraught with a Number of Methodological Flaws Worth Reviewing**S. Nyamhuno<sup>1</sup><sup>1</sup> USAID Southern Africa, Power Africa, Johannesburg, South Africa

Realistic evaluation is specie of theory based evaluation method that was pioneered by Ray Pawson and Nicky Tilley (1997). Primarily, realistic evaluation differs from the traditional quasi-experimental approach in that it aims at unpacking the black box by trying to precisely ascertain causality. Based on the epistemology of realism, realistic evaluation is centered on Context–Mechanism–Outcome (CMO) configuration. Context refers to the condition prevailing at the time of the intervention while mechanism refers underlying generative systems. Lastly, Outcome refers to the he product of the intervention. Without doubt explaining causality within an intervention is a remarkable achievement. So many program struggle to know whether M&E results are a contribution of the intervention or they can wholly attribute the results to their to the intervention. Traditionally, attribution needed a thoroughly defined counterfactual, which would eradicate any eternal influence. Realistic evaluation is created with attempting to explain certain results were attained in a program. However, there are flows in realism epistemology underpinnings. Realistic evaluation is not a method of evaluation per se, but a way of thinking around context, which might trigger certain outcomes. Data collection for CMO is very difficult to collect. In crime prevention program, for example getting data from offenders is usually hard. Where it is collected, there are always issues around its validity and let alone reliability. To ensure validity of the data, a lot of checks must be put in place to ensure data validity Contextual issues need a robust research in order for the evaluator to be certain that it is solely responsible for the outcome. Now external factors come into play and as it might be safe to suggest that only in experimental designs, particularly the randomized control trails (RCT) can only safely achieve this. A study by Gill and Turbin (1999) found out that the use of cameras by shop owners deterred a lot of stealing. On closer investigation it appeared that it was not the staff watching the cameras that helped, but cameras on their won where deterrent mechanism. Realistic evaluation seems to be limited to social aspects such as measuring crime reduction through certain mechanism in particular context. Other fields that are technical or which may involve sophisticated technological inventions and improvements are quite difficult to evaluate because of the inherent nature of realistic evaluation. It appears that realistic evaluation is suitable for the evaluation of micro projects and not macro projects. Policy makers, senior government officials, civil societies cannot use the realistic approach at a national level and worse still at an international level to measure the effect of an intervention. In any evaluation, one might chose a number of mechanisms that would operate in particular context, there are still limitations to that. Some of the mechanisms are quite difficult to define precisely as to come to a conclusive outcome. In some of the evaluations, it is practically impossible to cover all probable outcomes of mechanism. In another evaluation CCTV resulted in security workers relaxing, which led to increased theft.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 133 Handling Complexity: Approaches and Methods****O 387 - Revisiting Theory of Change of Interventions in Complex Systems***B. Douthwaite<sup>1</sup>, F. Ahmad<sup>2</sup>, G.M. Shah<sup>2</sup>, A. Mishra<sup>3</sup>*<sup>1</sup> *Boru-Consult, Enabling and Evaluating Innovation in Agriculture, Westport, Ireland*<sup>2</sup> *International Centre for Integrated Mountain Development, Strategic Planning- Monitoring and Evaluation, Kathmandu, Nepal*<sup>3</sup> *International Centre for Integrated Mountain Development, Livelihood Theme, Kathmandu, Nepal*

Increasingly, rural research and development programs realize they need to grapple with complexity in the systems in which they intervene. Complexity, through emergence, offers the chance of achieving much from small, well targeted interventions; conversely, complexity can make a mockery of linear planning such that large investments can yield little or no benefit. Harnessing complexity requires identifying emerging impact pathways early enough to respond. Developing and revisiting theories of change (ToC) helps programs to think about the impact pathways the program might catalyse and support, and to identify them from the start. However, while theoretically promising, little has been written about using ToC in this way in practice. This paper helps fill this gap by examining the case the International Centre for Mountain Development (ICIMOD) in Nepal that has institutionalized the use of ToC. The paper describes how Participatory Impact Pathways Analysis, developed in 2007, was introduced as part of a results-based management system. Key success factors included taking a staged approach, the use of ToC as part of results-based management, capacity development and consistent and high-level support for M&E to fulfil accountability and learning requirements. ICIMOD initially found revisiting ToC difficult because prospective ToC developed at the start of programmes was generally too general and aspirational to be testable. The paper describes a method ICIMOD developed for revisiting ToC based on identifying and describing retrospective ToC to explain early programme outcomes.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 133 Handling Complexity: Approaches and Methods

## O 388 - Deepening Participatory Practice to Build Inclusive Resilience; Where is the Evidence?

M. Apgar<sup>1</sup>, G. Lyn Higdon<sup>2</sup>

<sup>1</sup> Institute of Development Studies & ESRC STEPS Centre, Research Fellow in Participation, Brighton, United Kingdom

<sup>2</sup> Institute of Development Studies, Participation, Brighton, United Kingdom

In the current SDG era, development practice and consequently its evaluation is becoming more *challenge-driven*. We seek not just to meet targets, but to understand how we support change. We appreciate that change is non-linear, and that building adaptive capacity for resilience – and understanding trade-offs between SDGs – requires *complexity-aware approaches*. These are all steps in the pathway to achieving greater impact, and are beginning to influence evaluation design, push for learning oriented approaches that can support adaptive programme management. We see a conscientious rise in use of these approaches by bilateral donors, such as DFID, and by funders.

Yet there remains a need to go deeper, to engage with the overarching SDG tag line to 'leave no one behind' as we seek to achieve development outcomes. This requires that we engage with the *power dynamics* that lead to increasing inequality and exclusion of the most marginalised. If we care about the most excluded, then we should be asking ourselves, how do we build *their* resilience?

The downward accountability agenda of the development and humanitarian sectors is usefully pushing evaluation practice to listen more to the voices of those that programmes aim to serve – i.e. the spread of 'beneficiary feedback mechanisms'. More emancipatory approaches have been promoted and used historically to move beyond listed to the voices of the poor and marginalised and to engage them in the process. Some argue that when used well participatory methods create the potential for evaluation to become an empowerment process in its own right. There is experience of how to do this, my colleagues at IDS we have spent decades attempting it.

The challenge remains, however, that the complexity-aware (and adaptive management) rhetoric does not tend to overlap in practice with the participatory (in order to be emancipatory) rhetoric and practice. In this paper, I will present findings from an evidence review on the practice of **participation within complexity-aware and adaptive approaches to evaluation**. Using a simple participation spectrum we analysed documentation of complexity-aware MEL and conducted interviews with practitioners. I will show that participatory practice remains largely superficial and used to 'inform' programmatic decisions' at the field level, rather than engaging with and transforming power relations and influencing decisions further up the chain. The story, however, is not all negative, and I will share some examples of progressive practice that move beyond 'business and usual' where intentions are leading to deeper, and more emancipatory practice.

Friday, 5 October 2018  
13:45–15:15

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 134 Improving Science And Technology Policies And Programmes Through Evaluation****O 390 - Fostering Resilience through “Strategic” Monitoring and Evaluation: Evidence from the National Science Foundation Efforts to Build/Diversify the Scientific Workforce**C. Cosentino<sup>1</sup><sup>1</sup> *Mathematica Policy Research, Human Services, Washington DC, USA*

Access to critical information for decision making is the anchor of a resilient society. Government agencies and private donors rely on monitoring data systems (or data dashboards) to gain access to vital information about the use of grant funding and the production of expected outputs and outcomes. These systems are rarely used (effectively) for evaluation – as they generally fail to include the information needed for evaluation purposes. This clashes with growing information needs – driven in particular by the need to allocate limited resources among competing (and often changing) priorities. As monitoring data systems fall short in providing useful information for decision making, and evaluations often take too long to influence decisions in real time, attention has shifted to emerging technologies, such as machine learning, and crowdsourcing platforms, as potential solutions to these information gaps. This presentation will argue that monitoring data systems can be powerful tools for decision making – if strategically designed. It will showcase the value of leveraging the old and the new – that is, of designing data systems strategically to serve multiple functions (monitoring, research, and evaluation) and using them in intentional coordination with other sources of available information (such as data repositories or social media) and emerging technologies (such as machine learning).

Using two monitoring data systems from the United States National Science Foundation (NSF) – a longstanding one (created in the early 1990s) and a new one (developed in 2017–2018) – this presentation will seek to illustrate the ways in which these systems have evolved to meet the needs of different stakeholders. Specifically, the presentation will include findings from research based on the Louis Stokes Alliances for Minority Participation (LSAMP) Program data system to demonstrate the value and limitations of this early generation of monitoring data systems focused on cross-sectional measurement of participant characteristics and services. In contrast, the recently developed data system for the Research Experiences for Undergraduates (REU) Program (to be launched in Fall 2018) will help showcase a new generation of strategic data systems that seeks to advance two synergistic objectives. That is, it seeks to help NSF monitor program participation, education, and employment outcomes longitudinally, more effectively, and more efficiently – while also providing a data source for academics participating in the program and evaluators. Both of these programs – LSAMP and REU – are central to NSF's efforts to build a diverse, globally engaged, and well-prepared workforce in science, technology, engineering and mathematics that will foster the nation's resilience and competitiveness. More robust monitoring data systems can not only play a key role in supporting NSF efforts to meet this goal, but also support evaluation efforts that will provide critical evidence in a timely fashion to inform decision making in a rapidly changing social, political, and economic landscape. The presenter hopes to engage with the audience in a dialogue about the adaption of existing data systems – or creation of new ones – to serve “strategic” needs.

Friday, 5 October 2018  
13:45–15:15

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 134 Improving Science And Technology Policies And Programmes Through Evaluation

## O 391 - Comparing the Institutionalization Approaches of Korea and Netherlands Regarding the Integration of Foresight, Technology Assessment, and Evaluation

C.C. Chang<sup>1</sup>, T.Y. Liu<sup>1</sup>

<sup>1</sup> Science and Technology Policy Research and Information Center- National Applied Research Laboratories, Policy Research Division, Taipei, Taiwan

As indicated by several researchers, innovation system as a network among heterogeneous actors where the decisions are coordinated in the multi-level, or multi-actor arena concerning evaluation, technology assessment, and foresight. Micro-level interactions and linkages among individuals in the meso-level can yield their impact on the macro-level market economy and institutions comprising the consumer interactions, supply chain, and regulatory policies of the eco system (Elop, 2014; Valery, 2014; Lusch & Vargo, 2006).

Therefore, combined, the foresight, TA, and evaluation are employed in the increasingly complicated STI policies where the supply-side and demand-side policy instrument encompassing multiple agencies, three helixes, and end users are interconnected. As an east-Asian emerging democracy with technocracy legacy which has limited relational capacity, Korea is inclined to take a centralized, rationalist approach of integrating them.

In Korea, the institutional arrangement of integrating the foresight, TA, and evaluation has the following features:

- With a centralization approach in single national institute, KISTEP integrated foresight, technology assessment, and evaluation in the "Feasibility Assessment" with respect to large-scale R&D project to inform the program priority-setting based on the multi-criteria analysis.
- Affiliated with STI ministry, MSIP (Ministry of Science, ICT, and Future Planning), KISTEP gradually monopolized the mandates concerning the foresight, technology assessment, and evaluation for NSTC (Mikheeva, 2015; OECD, 2014; Chang, Ching-Chun, 2017).
- The foresight conducted by KISTEP predominantly focused on identifying the key technologies instead of on discursive articulation of societal needs with multiple stakeholders (Yim, 2011; Choi & Choi, 2016).

Assessing the reflexivity yielded by KISTEP's institutional context through comparison with regard to triple-helix structures, and TA institutionalization approaches, this paper identified the following primary weakness of KISTEP's institutional arrangement:

- There exists the contradiction between research evaluation and TA, which focuses on monitoring research community.
- Discouraging the inter-agency coordination and major STI policy transformation, KISTEP's affiliation with the STI agency, MSIP, undermined its capability to collaborate with mission-agency where the domain policies reside in.
- With focus on budget prioritization which works against the discursive reflexivity in the dialogues among multiple actors and sectors, it is challenging for KISTEP to serve as a countervailing force against the predestined STI budgets.
- Confined autonomy of evaluation and technology assessment under its institutional context, KISTEP exerted less autonomy and discretion regarding the STI program impact assessment, and professional expertises (OECD, 2014).

In contrast, Rathenau Institute of Netherlands take a totally different approach as opposed to KISTEP regarding the following dimensions:

- With focus on the science system assessment instead of specific research and development, Rathenau focuses on the system-level interactions and coordinations to provide reflexive strategic intelligence.

- Affiliated with science community along with other TA network institutes, Rathenau Institute have more autonomy and are more connected with the legislature, administration, civil society, and science community.
- Informing multiple mission agencies instead of single agency, TA network institutes can collectively exert autonomy against the administration, thereby broadening the scope of TA, foresight, and evaluation.
- Without focusing on politically irritating budget prioritization, Netherlands integrated them in the program level, aligning the R&D practices with real-time TA, foresight, evaluation to assume the autonomy of discursive dialogue to identify the unarticulated needs.

Friday, 5 October 2018  
13:45–15:15

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 134 Improving Science And Technology Policies And Programmes Through Evaluation**

**O 392 - Evaluating Effects of Public Support to Business R&D and the Economic Crisis: Spanish Evidence**

*A. Barajas<sup>1</sup>, E. Huergo<sup>2</sup>, L. Moreno<sup>2</sup>*

<sup>1</sup> CDTI, Studies and Communication, Madrid, Spain

<sup>2</sup> GRIPICO Group for Research in Productivity- Innovation and Competition, Department of Economic Analysis. Facultad de CC. Económicas y Empresariales. Universidad Complutense de Madrid, Madrid, Spain

Evaluation of R&D public policies is a key issue in any context, due to the expected effects of these interventions on many actors and aspects of the social and the economic environment. But it is even more important to evaluate these policies in a context of public expenditure restrictions, when budget allocation could be influenced by evidence based on accurate data.

As long as the role of public support for business R&D could be decisive mitigating the negative effect of private financial constraints during the crisis, the key question is whether this support is really having such an effect or, on the contrary, public aid is less effective in adverse economic conditions. From a different perspective the question to be answered is this: Should R&D public policies follow a pro-cyclical pattern (with decreasing budget in crisis periods), or, on the contrary is more appropriate for the society to implement a counter-cyclical model? Evaluation provides the rational and the methodology to address this relevant issue, although a strong and clear commitment of public authorities is required in terms of data availability to conduct a reliable research and a convincing evaluation.

The present study compares the effect of public support to business R&D on technological inputs and outputs before and during the recent economic crisis. To do so, we use information provided by the Centre for the Development of Industrial Technology (CDTI), which is the main public agency in Spain that grants financial aid of its own to companies for the execution of R&D projects. Specifically, we consider firms supported through CDTI programmes for periods the 2002–2005 and 2010–2012.

Impact evaluation is conducted using quantitative “matching” techniques. Our results suggest that, during the crisis, public support continued to have positive effects on the resources devoted to R&D activities, and also increased the technological outputs obtained from these resources. Therefore, based on this quantitative approach, counter-cyclical R&D policies would be recommended.

Friday, 5 October 2018  
13:45–15:15

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 135 Evaluation Challenges in Fragile Contexts****O 393 - The Role of Evaluation in Helping Build Resilience to Violent External Shocks: Lessons Learned and the Way Forward***K. Atanesyan<sup>1</sup>, S. Markova<sup>2</sup>*<sup>1</sup> *The World Bank Group, Independent Evaluation Group, Washington DC, USA*<sup>2</sup> *The World Bank Group, Public Opinion Research Group, Washington- DC, USA*

The last decade saw a drastic increase in violent conflicts affecting many countries around the world – the most tragic of them being the civil war in Syria and massive displacement of civilian population as a result. International community (including bilateral agencies and multilateral organizations, such as the UN system and the international financial institutions) responded quickly to the humanitarian aspects of the crisis. Many countries (particularly Turkey, Lebanon, and Jordan) opened their borders and provided food, shelter, and basic services to the refugees. International organizations and bilateral donors came up quickly disbursing financing and emergency assistance.

However, once the immediate emergency needs were addressed, all host countries – and particularly the communities most affected by the influx – started facing serious and protracted developmental problems, putting tremendous stress on the domestic systems for service delivery (health, education, municipal, etc.) and challenging their resilience. All stakeholders came to realize the primacy of these issues, and the importance of building up the resilience of domestic systems to withstand the crises and address possible political, economic, and social repercussions. Efforts in this direction rightfully (albeit somewhat slowly) became the focus of the international and domestic initiatives.

At the same time, several early evaluations of such efforts, conducted by various evaluation outfits of international and other organizations, including the World Bank Group, showed that many challenges persist in this area, including: (i) insufficient common understanding of the nature of the issues; (ii) ineffective division of labor among various actors (not based on respective comparative advantages) and frequent mission overlap and inefficiency; and (iii) general lack of forward-looking vision as to how strengthen the systemic resilience without disturbing at times fragile social cohesion.

As experience with dealing with these issues kept accumulating (including through evaluations), the potential critical role of evaluation as an important tool for ex-post and real-time learning in this regard became particularly prominent. The presentation will showcase experience with conducting evaluations in situations affected by violent conflicts and their aftermath, the role of various evaluative instruments and their impact on learning (for multiple stakeholders), accountability, and convening of international efforts. The presentation will synthesize experience from recent work done in evaluating humanitarian and development efforts in various contexts affected by violent conflicts and related forced displacement of population: Middle East (Syrian crisis), South-East Asia (Philippines/Mindanao, Indonesia/Aceh), Horn of Africa, other locations.

The presentation will also discuss the use and applicability of various evaluation techniques while conducting evaluations in situations affected by violent conflicts and suggest how to focus and use evaluations for informing various country stakeholders (governments, communities, development partners, civil society) and playing a convening role for putting together implementable medium- and long-term strategies to build up resilience of country systems. The presentation will cover qualitative and quantitative methods in evaluation used by the WBG's Independent Evaluation Group and Public Opinion Research Group, including case studies of select Country Program Evaluations (CPEs) and 2012–2017 Country Opinion Survey (COS) data from 27 fragile and conflict-affected countries.

Friday, 5 October 2018  
13:45–15:15

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### S 135 Evaluation Challenges in Fragile Contexts

## O 394 - “You Can’t Get There from Here”: A Program Evaluation in Iraq

*J. Kotun*<sup>1</sup>

<sup>1</sup> Independent Evaluator, Evaluation, Falls Church, USA

Evaluators working in conflict or post-conflict environments face unique technical, security and ethical challenges. Program funders and implementers are under pressure to achieve measurable positive results over short periods of time, working in a fluid chaotic environment. These realities further magnify real-world evaluation challenges. This presentation discusses an evaluation of FHI 360's *Foras* (“opportunities” in Arabic) job creation program in Iraq, funded by the United States Agency for International Development. Iraq's post-conflict environment challenges included the need to measure results of the short-term program that was designed in a bubble far removed from rapidly changing, chaotic real-world realities. The program implementers had only limited direct access to beneficiaries themselves, thus making the evaluation even more important as a rare opportunity to observe through outside critical eyes and ears. This presented a double-edged challenge for the evaluation team: the need for extra care in collecting adequate sample sizes to address the challenge of measuring small changes over the short term, while faced with severe security risks in accessing certain communities and groups of program beneficiaries. In addition, the Iraq jobs program was high profile and politically sensitive for the funder, stakeholders and beneficiaries.

This presentation will point to some of the knowledge gaps that are critical to recognize when working in conflict or recent post-conflict environments. Unfortunately, limited documentation of past real-world conflict experiences exists to guide evaluators. This reality is often exacerbated by the fact that many established, experienced evaluators eschew working in these environments. While this presents opportunities for less experienced evaluators to accept the challenge, it also requires a steep learning curve with many potential blind spots. Of particular interest to other evaluators working in similar environments, the presentation will focus on how the evaluation's design and team management approach rapidly adapted through trial and error in the early stages to identify blind spots, improvise ways to access program beneficiaries in insecure areas, modify sampling and blend methods to meet the real-world challenges, including a broad range of beneficiaries such as internal displaced persons (IDPs) in camps, university students, highly educated professionals, and casual laborers. Other challenges included an inexperienced local evaluation team, language challenges of conducting focus group discussions and interviews in Arabic while reporting in English, and severe constraints on access to the field due to ongoing security concerns. The presentation will discuss the team management approach to training and coaching the local evaluators to become sensitive and observant “eyes and ears”, overcoming inter-cultural and language challenges to data quality through a highly collaborative and participatory approach to evaluation team management working in Iraq's post-conflict reality. The discussion will also include how the evaluation team addressed the challenges of working in Iraq, what worked, what didn't work, what was important, and what turned out to be less important and lessons learned for program evaluations in conflict environments including work-around solutions to technology, security and logistical challenges, understanding the underlying political landscape, and managing expectations.

Friday, 5 October 2018  
13:45–15:15

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**S 135 Evaluation Challenges in Fragile Contexts**

**O 395 - Making Resilience Visible: Results from a Rigorous Impact Evaluation of a Land-Use Planning Intervention in the Philippines**

*G. Leppert<sup>1</sup>, M. Lech<sup>1</sup>*

<sup>1</sup> DEval - German Institute for Development Evaluation, Competence Centre for Evaluation Methodology, Bonn, Germany

In this presentation, the authors approach the question of how evaluation can be used to uncover resilience among households and communities. They exemplify their approach and findings with examples from a rigorous impact evaluation of a comprehensive land-use planning intervention in the Philippines.

The complex intervention was implemented between 2006 and 2015 by the Philippine-German cooperation and national agencies. It supported municipalities to conduct enhanced land-use planning, including core aspects of environmental sustainability, disaster risk management, local governance and socio-economic development.

The comprehensive approach aimed at reducing vulnerability to negative effects of uncontrolled development and to multiple hazards, including human-made risks, and climate change. For the first time, the intervention applied a holistic planning approach encompassing all municipal ecosystems, "from ridge-to-reef", including other municipalities in the same "watershed". The intervention outcomes and impacts aimed at pro-actively improving resilience of municipal households in the field of welfare, disaster risk management, climate change, and environmental conservation.

The authors show the measured impacts of the intervention on municipal planning capacities and on plan quality and comprehensiveness. The authors present the impacts on several indicators related to vulnerability and resilience in aforementioned impact fields and shed light on the implications resulting from mainstreaming the approach into national policy-making.

Especially the results in the field of disaster risk management proved interesting, which is of vital importance for the Philippine context and the vulnerable location of the study region. Hence, the authors added a secondary study to further elaborate on the findings in the field of disaster risk management. To back up results from the impact assessment, the team utilized high-resolution remote sensing images to detect visible changes in disaster recovery and preparedness in selected municipalities. Using machine-learning based land-use classification techniques, the team is able to evaluate the paths of post disaster reconstruction and resilience.

Beyond the presentation of evaluation results, the evaluation team demonstrates the benefit of geospatial and remote sensing data to approach resilience and environmental change complex evaluations and mixed-method evaluation designs. They show how these methods complement traditional evaluation methodologies in all phases of impact evaluation.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 136 Evaluating Civil Society Organizations

## O 396 - Organisational Reviews of Civil Society Organisations: How Can Usefulness Improve While Costs Are Reduced?

V. Hildenwall<sup>1</sup>, A. Millard<sup>2</sup>

<sup>1</sup> FCG Sweden, M&E Unit, Stockholm, Sweden

<sup>2</sup> Policy Research Institute, None, Chamoson, Switzerland

Bilateral and multilateral donors work with civil society organisations (CSOs) – organisations that include membership-based, cause-based and service-oriented groups – because CSOs are key actors in development. CSOs implement aid projects and programmes in developing countries, both on behalf of donors and in their own right. As thematic/local context specialists with ability to reach out to end beneficiaries, they have clear comparative advantages in certain areas. They are also important contributors to donor development co-operation policies and as advocates for development issues.

**Rationale:** When selecting CSOs for collaboration, a donor need to ensure that the organisation can contribute to achieving the donor's overall development objective(s) in an efficient and meaningful way. A common way of doing this is to let the organisation undergo an **organisational review**. This is (simply put) an exercise with the purpose to determine to what extent an organisation is fit for purpose: does it share the same basic values? Are there accurate internal financial control systems in place? Does the organisational structure (incl. staff management) seem appropriate? Does the organisation achieve results?

**Organisational reviews form part of the basis for donors' funding decisions. It is thus imperative that these reviews are reliable and provide correct information. At the same time, they are often conducted with resource constraints – both in terms of time and financially.**

**Objectives sought:** Having implemented a number of organisational reviews for Sida (Sweden) and Norad (Norway), some issues for concern has arisen that we would like to discuss through this paper. From our perspective as independent consultants, we can see that both sides – the donor and the CSOs – struggle through the collaboration process: while the donor may experience high transaction costs for assessing and coordinating a great number of organisations, the CSOs tend to see the conditions set by the donor as extremely demanding. Despite the fact that an organisational review may be an informative experience for the organisation, it is also highly time consuming – staff members need to make themselves available for interviews, relevant documentation has to be provided by short notice, there might be expectations to read and comment upon draft reports, etc.

This leads us to pose the following questions:

- What lessons can we draw from our experience of conducting organisational reviews for two different donors? Where did they differ?
- How do we ensure that organisational reviews focus on the most relevant issues?
- Do the reviews need to be donor specific? What is the experience of co-commissioning organisational reviews? Advantages/disadvantages?
- How can donors and CSOs work together for achieving a more efficient collaboration?
- How can reviews be more useful to the CSO under review?

**Brief narrative and justification:** CSO play an active and important role in development co-operation. Can donors' organisational reviews of CSOs be made more efficiently? Evaluation practitioners and commissioners should be interested in discussing new ideas and innovative approaches for organisational reviews – an important tool for donors to determine if and how to collaborate with a CSO.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 136 Evaluating Civil Society Organizations

## O 397 - Relevancy Evaluation in Cooperatives' Performance. A Guiding Material for Cooperatives' Performance Evaluation

A. Yasin<sup>1</sup>

<sup>1</sup> ESDC, M&E Department, Ramallah, Palestine- State of

The cooperative sector makes a significant contribution to the socioeconomic development in Palestine. Cooperative contribute to job creation, income generation and poverty reduction, As member-based social enterprise, the cooperative organization contribute to the economic change, while maintaining fair economic participation of their members, and practicing democratic management. Additionally cooperatives have their social responsibilities to their societies and contribute to strengthening the civil society concept.

Co-operative principles, and values, as defined by the International Co-operative Alliance, have served as the guiding, as well as defining, tenets of co-operative founding and functioning. Co-operatives that have stayed true to these tenets have delivered immense value to their members, as well as their communities and customers. In this paper, I tease out the relevance of these tenets.

However, the recent studies still highlighting a key problem of cooperatives movement in Palestine remains *the lack of right understanding of coops principle and lack cooperative model*; the key stakeholders believe that there is crucial need to support development of model co-operative in order to use these successful models as a mean for sharing and spreading the right cooperatives concept and modality among the other cooperatives.

The relevance in cooperatives can be studied through analyzing the cooperatives generic aspect; as a) cooperatives case or the justification of the cooperative; b) the capital and investments; c) the membership file; d) work systems and bylaws; e) shared needs and; f) services and products. Other field of study is relevance in cooperative's concept and international cooperatives principles, also with cooperatives values.

Relevance analyses can be used in pre registration phase as tool of verification by the registration authorities, this will help in starting a active cooperatives with right cooperative case. Relevance analyses can be used for cooperatives strategic area as vision, mission, and strategic objective. Good governance practices is another field of analyses in the relevance analyses. The relevance analyses can also depend on the level of cooperative's maturity and development. The business case and production capacities is also an important field of the relevance analyses.

Relevance analyses is a comprehensive approach that can be used at different levels of cooperative development with the main output of this approach is a diagnostic report that is used to define the field of improvement within the cooperative. In this context, I have developed a manual that contains important and helpful elements that guide the Palestinian cooperative on how they can become model cooperatives that match their activities with members and communities needs and priorities. This manual is the first product of a continuous learning and development process.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 136 Evaluating Civil Society Organizations**

**O 398 - Evaluating the Work of German 'Stiftungen' Through Bonding, Bridging and Linking Social Capital: A Case-Study on Kyrgyzstan**

C. Pierobon<sup>1</sup>

<sup>1</sup> Bielefeld University, Post-Doctoral Fellow, Bielefeld, Germany

Over the years, German political foundations (*politische Stiftungen* / *parteinahe Stiftungen*) have been assigned a role of pivotal importance in the development of democracy and the strengthening of civil society inside and outside of Germany. Between 2005 and 2014, the federal budget allocated to them registered an increase of about 50%, reaching the amount of 466 million Euro in 2014. The increase of the budget available has gone hand in hand with the increasingly important role assigned to the *parteinahe Stiftungen* as foreign policy tools.

Although German political foundations are among the oldest, most experienced and biggest actors in international democracy assistance, the literature available on their international work is very limited. The paper addresses this gap in the literature by providing an empirical analysis of the work of the Hanns-Seidel-Foundation (HSS) and the Rosa-Luxemburg-Foundation (RLS) in Central Asia. More precisely, it deepens our understanding of the role and contribution(s) of these peculiar transnational actors as external providers of democracy aid by looking at the social capital produced. Several studies have provided evidence of the impact of social capital on better governance and more effective policies and an incremental acceptance of the concept of social capital into decision-making processes that are aimed at achieving development objectives has been registered. In the analysis, three types of social capital – bonding, bridging and linking are taken into account.

The contribution of this study is twofold. On the one hand, the article provides new insights on the work of German political foundations in Central Asia, a field that is still unexplored. In particular, by offering a micro-level analysis of the engagement of the HSS and RLS in Kyrgyzstan, it enhances our knowledge on the functioning of these peculiar providers of democracy aid on the ground. On the other hand, the article tests the use of social capital and its three types – bonding, bridging, and linking – as heuristic tool for the examination of the democracy assistance provided by transnational actors. In so doing, it sheds light on the potential of this concept not only for the evaluation but also for the conceptualization of more effective programs aimed at strengthening citizen participation.

Friday, 5 October 2018  
13:45–15:15

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 137 The Role of Evaluation in Promoting Support for Development Aid**

**O 400 - Multi-Stakeholder Initiatives – the Case of EITI: Paving Pathways to Impacts through Evaluations**

*L. Neumann<sup>1</sup>*

<sup>1</sup> *LBN Strategies, LBN Strategies, Koeln, Germany*

In 2003, as a result of the public debates of the role of multinational extractive companies operating in developing countries in Africa, Asia and Latin America and research pointing to phenomenon that became known as the 'resource curse' and 'paradox of plenty', the Extractive Industries Transparency Initiative (EITI) was founded. This multi-stakeholder initiative is nowadays implemented by 51 countries worldwide (e.g. Afghanistan, Germany, Liberia, Kazakhstan, Nigeria, Norway, Peru, Philippines, Ukraine) and widely considered as a global standard to promote the open and accountable management of oil, gas and mineral resources. Any multi-stakeholder initiative – including the EITI – can be considered as a response to a governance deficiency in a country or a sector. More specifically, as the response to a non- or dysfunction of state institutions and of their relation to other relevant stakeholders.

The way in which a country is implementing a multi-stakeholder initiative, or even the reasons why it is implementing it, can be expected to change as the multi-stakeholder dialogue deepens, so the M&E framework needs to be able to accommodate that dynamic process. The evaluator's key challenge when choosing the approach is to find a flexible, yet robust framework to allow identification and attribution of causal effects, which – in foresight of the further evolution of the initiative – should not squeeze too narrowly or freeze the dynamics. A multi-stakeholder initiative is not a project investment with a fixed log frame.

As the findings of the author's assessment of the effectiveness and impact of the EITI have shown, based on a reconstruction of a generic results model and contribution analysis applying a mix of methods of empirical social research (quantitative and qualitative data collection methods and data analysis) including a perception-based approach by survey, evaluation can be conducted in a way that is robust enough to survive short-term changes, but also flexible enough to measure outcomes and impacts over the long term. Whereas mere quantitative analyses of panel or cross-sectional data to tease out statistical relationships driving observed changes by regression analysis were found not to be indicative to address key issues of impact and causality. In four key arenas such as fiscal transparency, public debate, anti-corruption, and trade and investment climate by means of evaluation outcomes can be attributed, as the result of the initiative's activities. The paper promotes a certain openness in order to uncover impacts that have not been anticipated by anyone; nor directly stated, categorized or measured, because they are the kinds of effects that might suggest much longer-term trends that have yet to become visible enough to be more accurately captured and measured. In those circumstances a participatory evaluation design is key to any meaningful evaluation.

Key questions center on the following: How should evaluation of multi-stakeholder initiatives be designed to support implementing countries in their resilience and their momentum for change. Did the multi-stakeholder initiative increase the capacities of stakeholders to participate in informed decision-making for public or corporate policy?

Friday, 5 October 2018  
13:45–15:15

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**S 137 The Role of Evaluation in Promoting Support for Development Aid****O 401 - Engaging The Private Sector for Agricultural Development – an Evaluation of Germany's Policy Approach***M. Kaplan<sup>1</sup>, N. Herforth<sup>1</sup>, S. Brüntrup-Seidemann<sup>1</sup>*<sup>1</sup> *German Institute for Development Evaluation, Sustainable Economic and Social Development, Bonn, Germany*

The SDGs emphasize the need for global partnerships between different actors to foster sustainable development. The private sector has become an increasingly important partner for development cooperation (DC) both internationally but also in Germany that can contribute to promoting the agricultural sector in developing countries. Especially the “Marshall Plan with Africa” by the German Federal Ministry for Economic Cooperation and Development (BMZ) emphasizes the role of the private sector. However, critics particularly from civil society organizations doubt the positive effects of this approach for smallholder farmers in rural areas as one of the main target groups of German DC. Furthermore, they suspect that cooperating with the private sector as a still rather new player in DC, who is not naturally committed to the goals of DC, might even result in human rights violations.

There are many programmes, in which German Technical Cooperation attempts to collaborate with German, international, and local companies, either cross-sectoral or exclusively in the agricultural sector. This evaluation analyses the extent to which the German portfolio, i.e. the strategies and programmes, is suitable to contribute to the goals of German development cooperation in the agricultural sector. It particularly worked out the concrete benefits of the private sector as a newer player in DC. Taking into account criticism of this approach from civil society, the evaluation also analysed the mechanisms of German DC for ensuring that activities with the private sector respect international human rights standards and principles. We carried out content analyses of strategic and programme documents and reconstructed the theory of change for cooperating with the private sector in agriculture in German Technical Cooperation. These data collection methods were accompanied by 78 qualitative interviews with different stakeholders of German DC, both in the ministry and in implementing agencies, companies cooperating with German DC, and civil society organisations. We triangulated the data sources to capture the different perspectives of the highly diverse stakeholder groups to identify benefits, but also challenges and areas of conflict with regard to this approach, and to increase the validity of our results.

The evaluation found that cooperating with the private sector is relevant and in general suitable to stimulate economic growth and through this pathway may contribute to poverty reduction and food security as two of the main goals of German DC in the agricultural sector. However, the evaluation also demonstrates that strategies and programmes of German DC do not clearly identify and describe the concrete developmental benefits of engaging the private sector. These conceptual limitations have led to uncertainty about potentials, usefulness and goals of this approach. Furthermore, the German approach does not translate the idea of a partnership between DC and the private sector into practice, as there are significant discrepancies between these partners in terms of goals, procedures, and common understanding that cause high transaction costs with regard to their cooperation. Finally, German DC lacks adequate mechanisms to ensure the ex-ante assessment and monitoring of human rights-related risks.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 138 The Transformative Power of Evaluation**

**O 403 - Exploring Methodologies to Measure Power Transformation in Social Fabrics**

*E. Emam*<sup>1</sup>

<sup>1</sup> *Institute for War and Peace Reporting, Programme, Tunis, Tunisia*

Over the last ten years, power analysis has been introduced as a tool that generates deep understanding of the context and navigation between options in order to address complexity. Following cumulative experiences with power analysis to address social justice, it was given more attention by development workers as a lens through which the roots and causes of today's inequality and injustice are better visualized and accordingly more efficient strategies can be drawn. Issues of social justice, equity, and poverty reduction are now viewed in a globalized context, which adds more complexity for programming and interventions. Hence viewing those issues of vulnerability should involve innovative analyses to design a realistic and significant approach to address them; otherwise they are not sufficiently spelled out. And "without a solid understanding and analysis of a particular situation that development cooperation sets out to modify, there is a real risk of failure" (Pettit 2013: 3). Many sociologists, political scientists and anthropologists have discussed power but little of these debates appear to have percolated into development studies (Mosedale 2013). Remarkably, there are still more work should be done in terms of understanding the concept of power in the developmental context. This paper is an attempt to revisit power analysis frameworks as tools to plan, design and evaluate change in power structures with focus on conflict affected regions. It will try to answer or probably reflect on the following 2 questions: 1- Why power analysis is important in today's development programs and how it can improve the aid effectiveness? 2- How power tools can present a holistic approach to evaluate development and resilience programs? Furthermore, practical experiences on using power cube as a tool to measure social cohesion programs in Lebanon and Jordan will be introduced in the paper.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 138 The Transformative Power of Evaluation

## O 404 - From Impact Heterogeneity to Implementation Heterogeneity: Equity Measurements for Informed Programming

*D. Chandurkar*<sup>1</sup>

<sup>1</sup> Sambodhi Research & Communications, Vice-president, Noida, India

More often than not, evaluation of equity effects of any program focuses on disaggregation of key outcomes with respect to various population strata-caste, wealth, religion, gender. The idea being-description of outcome differentials and inequities across these determinants. While this measurement of impact heterogeneity is important, deeper insights are necessary for understanding interplay of these equity determinants in identifying the most marginalized. Better programming can then be done to address inequities and improved targeting of the most-marginalized. Thus, it becomes important for evaluation to not only facilitate a comprehensive understanding of equity effects across multiple dimensions; but also to inform programs of the most-neediest such that strategies could be developed for improved targeting and addressing 'gaps'.

With these considerations, role of the evaluator thus expands beyond the conventional measurement of inequities and equity effects-to that of an active agent aiding programmes in solving for inequities. The evaluator's role thus becomes 'developmental'-facilitating programmes in addressing the needs of different groups differently-initiating implementation heterogeneity. Based on the experiences of a large-scale complex health systems intervention in India, the presentation would share experiences and lessons learnt from the four-stage process of **Definition-Intention-Measurement and Programing**-adopted for informing actions addressing health inequities.

In the first stage of the process-Definition-equity in the intervention context is defined. In the health system context-it was defined as differential access to services, utilization of services and differential outcomes. Along with classical equity determinants-caste, wealth, religion, contextual determinants-maternal education, maternal age and birth order are delineated.

This is followed by the second stage of Intention in which program intentions to address equity are described. Implementation pathways detailing how the specific interventions are going to address equity and the mechanisms therein are specified. Realist evaluation approach for context, mechanisms and equity outcomes are also explored.

In the third stage of Measurement, multiple methods and analytical techniques are employed for comprehensive measurement of equity aspects and getting deeper insights into intersectionality issues. The methods being-having adequate sample size for disaggregated analyses by equity stratifiers, aggregating stratifiers for defining marginalization, interaction analysis for equity determinants, multi-level modelling for describing effects of equity stratifiers including geographical inequities and Classification And Regression Trees (CART) for understanding intersectionality. Measurement then informs the fourth stage-Programing-facilitating design and implementation of appropriate intervention sets for better addressing equity and delivering equity outcomes.

Lessons learnt from the experience suggest the need for context consciousness on behalf of the evaluators in appreciating equity issues. More often than not, equity is assumed as an outcome with specifics of ways and means to address equity not detailed. Evaluators thus need to inquire of these specific ways in which the equity issues are being addressed. The challenge for evaluators is of not being limited to assessing impact heterogeneity, but of going beyond the conventional mandate and assuring heterogeneity in implementation. The onus on evaluators is therefore to keep the equity conversations going and help the program focus and refocus around equity.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 139 Evaluating the Impact of Research Infrastructure**

**O 405 - Evaluation Methods Old and New! The Use of Bibliometrics and Expert Panel Review to Better Evaluate Scientific Impact**

A. Ventimiglia<sup>1</sup>, D. Campbell<sup>2</sup>, E. Bailly<sup>1</sup>

<sup>1</sup> Science-Metrix, Evaluation, Montreal, Canada

<sup>2</sup> Science-Metrix, Bibliometrics, Montreal, Canada

Resilient societies are those that can incorporate many different approaches to find solutions to societal challenges. In line with this idea of embracing varying approaches, and in line with the conference strand, "rethinking evaluation methods and methodologies." Science-Metrix will present a novel mixed-methods approach we used in an evaluation of a high-risk and ambitious basic research funding program in the biological sciences. This evaluation focused on assessing the extent to which the funded research projects generated major advances in a sub-field of biology.

The evaluation combined traditional methods (e.g. document and administrative data reviews, e-survey, interviews, site visits) with less conventional methods (e.g. bibliometrics and expert panel assessment) such that the latter set of methods filled data gaps and helped triangulate findings extracted from the former. There was immense value in this approach because the unconventional methods bring with them respectively a strong quantitative element and a robust assessment of the science, complementing and validating the rich data gleaned from documents, survey, interviews etc. Taken all together, this allowed our evaluation to paint a detailed portrait of the research achievements and to make specific conclusions and recommendations that would not have been possible without the mixed method approach. For instance, the bibliometric analyses provided a comprehensive assessment of the volume and citation impact of peer reviewed papers produced by the researchers, and the expert panel was able to identify the best examples of scientific achievements that were aligned with the program's goals.

An additional innovative aspect of this project was the fact that the bibliometric method used a big data approach. In addition to traditional bibliometric indicators of publication citations, a search through the full text of roughly 200,000 open access publications was performed to find instances of uptake of research project outcomes that were not publications (e.g. web-platforms).

The presentation will begin by outlining one of the evaluation questions we were mandated to answer in the context of this evaluation, and will then take the audience through a discussion of how each method contributed to answering that question, leading to findings that were supported by multiple methods and finally resulting in well-rounded conclusions and recommendations. Finally, we wish to mention that the lessons we have learned from this project were specifically in the context of examining impacts in the research funding or scientific sphere. That said, in the interest of inclusivity and learning, we will encourage dialogue at the end of our presentation, on how our specific mixed-method approach might be utilized in a social science or other setting.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 139 Evaluating the Impact of Research Infrastructure****O 406 - Ex-Ante Evaluation of Research Infrastructures: Benefits, Methods and Approaches, Steps in Implementation, and Other***I. Spanache*<sup>1</sup><sup>1</sup> UEFISCDI - Executive Agency for Higher Education, Research Development and Innovation Funding, Bucharest, Romania

Even though a Research Infrastructure (RI) is initially built for scientific reasons, its impacts can spillover to many other different fields, such as the economic activity in the region, human resources, the socio-cultural life etc. And, at ex-ante level, evaluation can lead to getting the most out of RIs, to an increased multiplying effect of impacts, improving implementation and ensuring accountability. In this context, the paper will give arguments regarding the added value that can be brought by evaluating and monitoring RIs and why policy-makers, RI managers and public authorities should consider it. And as there is no "one-size-fits-all" solution, it will explore the way in which different evaluation methods and approaches (with a focus on ex-ante evaluation) can be adapted for evaluating RIs, and propose a sequence of steps that can be followed for ensuring a sound monitoring and evaluation process in accordance with the life-cycle of the RIs.

When it comes to the evaluation of Research Infrastructures, the existing literature is not as rich as in other fields. Certain guiding documents and a couple of articles have been published by different organizations (e.g. OECD), and by individual authors, that focus both on evaluating individual RIs and the entire landscape of RIs at national or European levels, as part of elaborating RI Roadmaps. However, these documents cover only certain elements of the evaluation and monitoring process and, sometimes, from the perspective of specific evaluation methods, such as CBA, foresight enriched methods etc. And only a few RI evaluation reports are available online entirely. In this context, the paper will have a more comprehensive approach, by trying to correlate the RI lifecycle with stages in the extended process of monitoring and evaluation. Subjects that will be covered by the paper are: Why do ex-ante evaluation of RIs? – functions of evaluation; What's in it for me? – arguments that can be given by evaluators to policy-makers, RI managers, funders of RIs, RI staff members and other actors, regarding its benefits; an overview of the RI lifecycle correlated with stages in the extended process of monitoring and evaluation; methods and approaches (e.g. theory of change, counterfactual impact evaluation, CBA, how and why to do a baseline study, etc.) coupled with explanations and examples that show their applicability in the case of RIs; a list of to do steps that should be taken into consideration; examples of indicators and evaluation questions; and other. Especially in certain countries from Eastern Europe, but not only, evaluation in general and evaluation of RIs, in particular, is underdeveloped and looked down on. And this paper will not only serve as a short guide for doing ex-ante evaluation of RIs, but it will also provide a solid argument basis for evaluators and other actors to advocate for undergoing evaluations of Research Infrastructures. And the challenge becomes even greater in the context of the fact that RI managers and the representatives of the scientific community tend to consider as important only the scientific impacts and disregard other types of impacts (socio-economic).

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 139 Evaluating the Impact of Research Infrastructure

## O 407 - The Socioeconomic Impact Assessment of Research Infrastructures: A Critical Review of Existing Approaches

*F. Giffoni*<sup>1</sup>, *H. Kroll*<sup>2</sup>, *A. Zenker*<sup>2</sup>

<sup>1</sup> CSIL Centre for Industrial Studies, Development and evaluation unit, Milan, Italy

<sup>2</sup> Fraunhofer Institute for Systems and Innovation Research ISI, Competence Center Policy – Industry – Innovation, Karlsruhe, Germany

Since the 2000s, a new generation of Research Infrastructures (RIs) has been actively developed all over the world, in all fields of science and technology to provide scientific support for the resolution of major societal and environmental challenges. Moreover, there is an increasing international competition in hosting facilities at the frontier of scientific and technological knowledge to promote economic competitiveness.

The question about the socioeconomic impact of science arises given the increasing amount of public money invested for research facilities at a time of tightening budget constraints. In a situation of competing claims for budget, Government and research funding institutions come under increasing pressure to justify their investments by demonstrating the value added that RIs provide. Moreover, there is consensus among most EU and OECD Member States on the need to promote evidence-based strategies for coordinated investment in RIs and to closely link them to evaluations and impact assessments. All these elements come together in reviving a strong interest and demand for methods for evaluating the socioeconomic impact of RIs.

While it is thus clear that impact assessments play an increasing role in decision-making processes on RI investments, the evaluation models that currently used remain heterogeneous in both objective and conceptual foundation. To understand how RIs develop impact on economy and society, we critically review the existing theoretical literature as well as the most common evaluation approaches, outlining the conceptual approaches and criteria used to estimate the socioeconomic contribution of RIs. For each approach, we highlight key assumptions, objectives, advantages and methodological limitations.

We group relevant studies identified from the literature in five approaches on the basis of the theoretical framework and methods used to value the benefits from RIs. These approaches are: 1) Macroeconomic Approaches (Demand-based input-output analysis and Computable General Equilibrium models); 2) Methodologies grounded on the Knowledge Production Function Approach; 3) Microeconomic Approaches (i.e. Cost-Benefit Analysis); 4) Approaches identifying RI impact at the firm level and 5) Mixed Method Approaches, including Case-Studies.

In summary, the proposed review contributes to the debate by a critical review of existing practices, common patterns and by offering a comprehensive overview of methodologies and data requirements for the estimation of the socioeconomic impact of investments in RI. It will address the increasing demand for a 'broader impact assessment' of publicly funded research facilities that is being articulated in many policy and stakeholder circles. Concluding, it will outline a first sketch of foundations for an integrated methodology for impact assessment that could integrate and reconcile relevant elements from different approaches.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 140 New Developments in Monitoring for Evaluation****O 408 - From Theory-based Evaluations to Theory-based Monitoring-and-Evaluations?**P. Lefebvre<sup>1</sup>, L. Virginie<sup>2</sup><sup>1</sup> Mines ParisTech PSL Research University, i3-CGS Interdisciplinary Institute for Innovation- UMR CNRS 9217, Paris, France<sup>2</sup> IWEPS, Research & Evaluation, Belgrade, Belgium

From theory-based evaluations  
to theory-based monitoring-and-evaluations?

The opportunity to mix evaluation and monitoring, often evoked, has been sometimes thoroughly discussed. Among publications, positions on this hypothetical mix may vary.

On the one hand, some would stress differences and discontinuity between evaluation and monitoring (SFE, 2009) while other publications put forward the idea that monitoring is a way to regularly collect information, so that it can be very useful at the time for evaluation (Prennushi & al., 2001 ; Kusek & al., 2004).

Our intent in this paper is to explore the idea that monitoring may be highly useful for evaluation – but that this usefulness is all but automatic. Such a usefulness appears under two successive conditions : *first*, the policy considered entails long and complex chains of causality and the aim of evaluation is foremost to get a better understanding of these chains, so that a theory-driven evaluation is particularly suited ; in such a case, 'theory-based monitoring systems' may deeply enhance evaluation. *Second*, the monitoring system is designed not (only) for monitoring purposes but explicitly for evaluation purposes and in coherence with the evaluation framework. This is 'monitoring-for-evaluation' as distinct from plain or broader 'monitoring'.

In a first step, the paper details the four potential benefits of monitoring from an evaluation perspective. Briefly sketched, they are: promoting, cognitive, normative and instrumental. Theory-based monitoring *increases the probability* to achieve an (theory-based) evaluation of a policy. With its systematic set of indicators, theory-based monitoring helps enhancing the *knowledge* function of evaluation (more time for qualitative interviews, times series data, test of hypothetical causal links of the program theory, etc.). It also serves as a basis for an evaluative *judgement* of the relevance, coherence and effectiveness of the policy. Last, theory-based monitoring is *instrumental* in so far as it helps to diffuse and share, from the upstart of the policy and for years, a representation of the intervention logic so that, when it comes to decision after evaluation, different stakeholders are deeply aware of the consistency or not between scenarios of decision and evaluation, so that they can (re)act to increase such a consistency.

In a second step, drawing on two innovation cluster's policies set up in 2005 and 2006 in France and Wallonia and on their three phases, three evaluation exercises (one between each phase plus one more recent; six evaluations as a total) and their evolving monitoring systems, it appears that the empirical evaluations and the counterfactual evaluations achieved hardly relied on monitoring data. Surprisingly, the same is roughly true for the only theory-based evaluation conducted. We show that the corresponding monitoring system had neither been designed for evaluation purposes nor in a theory-based perspective. We then explicit how, drawing on a governmental wish, we designed at the same time the guidelines of a future theory-based evaluation of the 3<sup>rd</sup> phase of the Walloon policy and a set of logically related monitoring indicators, dedicated to monitoring and evaluation purposes.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### S 140 New Developments in Monitoring for Evaluation

## O 409 - Graduate Tracking System as a Measure of Evaluation of the Educational System

*J. Dąbrowska-Resiak*<sup>1</sup>

<sup>1</sup> Foundation for the Development of the Education System, Research and Analysis Unit, Warsaw, Poland

According to the Eurydice Report (2014): “Modernisation of Higher education in Europe: Access, Retention and Employability” almost all European countries have already developed their own graduate tracking systems in order to monitor graduates’ professional development and their career. Also individual HEIs conduct their own graduate tracer studies. However, these systems use different research methods, use different data and present different results. Some of them use administrative data, other develop their own questionnaires and some are sporadic and ad-hoc studies. They also serve different purposes: forming part of evidence-based policy; assessing the relevance of study programmes to the labour market demand; researching the employability of graduates and factors influencing it; giving feedback to HEIs and students. Their thematic area concerns mostly the school-to-work transition, but it refers also to educational policy, to the labour market changes and to youth policy.

In 2016 European Commission adopted the New Skills Agenda for Europe, aiming, among others, to improve information and understanding of trends and patterns in demands for skills and jobs, with Graduate Tracking as one of the key actions to undertake in the future. However, this proposition is still underway.

In Poland, various studies are currently being conducted on the career paths of graduates, both central and decentralized (conducted by HEIs or research institutes). One of the most important is Polish Graduate Tracking System (<http://ela.nauka.gov.pl/en/>) implemented by the Ministry of Science and Higher Education and developed by the University of Warsaw (PEJK, [www.pejk.uw.edu.pl](http://www.pejk.uw.edu.pl)). In 2017 the Foundation for the Development of the Education System (FDES), National Agency for the Erasmus+ program launched independently a tracer study concerning professional and educational career of HEI graduates who participated in mobility under Erasmus program during their studies.

The proposed paper discusses different methodologies and different approaches towards different graduate tracking systems in Europe. It focuses mainly on Polish examples of Polish Graduate Tracking System and graduate tracer study by FDES in comparison to other graduate tracking systems, e.g. conducted by Cereq in France or Longitudinal Study of Young People in England (LSYPE). The main aim of the paper is to find the answer whether these analysed graduate tracking studies are the relevant measure to evaluate the educational system, the study programmes or the relevance of the studies to the labour market demand.

The key questions are as follows:

- How do different graduate tracking systems help evaluate the educational system? In what aspects can their results be helpful?
- What is the impact of such studies on the key actors on the labour market, on the education system, and on individuals?
- How do results of these studies vary with different methodologies used and different thematic areas? Is the use of administrative data helpful in terms of quality of results?
- What research scheme would be most suitable and relevant in different situations?

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 141 Documenting Program Process and Implementation****O 410 - An Innovative Evaluation Approach Towards High Quality and Rigorous Evidence: Documentation of Program Processes***F. Nyangara*<sup>1</sup><sup>1</sup> ICF, International Health and Development IHD Division, Rockville- Maryland, USA

**Rationale:** The global development and social programs designed to reduce poverty and malnutrition among vulnerable populations are complex, multi-faceted, and multi-sectoral in nature. It is also widely acknowledged that evaluations of these programs provide great opportunities for producing robust evidence and learning across projects and development contexts to guide ongoing improvement and further investments for scaling up. However, most of such evaluations are conducted as a matter of formality (accountability) and without a clear and purposive learning agenda. The evaluations have often relied on a combination of perspectives of a few stakeholders through interviews and sub-optimal monitoring and evaluation (M&E) data to understand how the program worked and its achievements retrospectively. Consequently, decision makers have learned very little from these evaluation efforts. Hence there have been calls for intentional efforts to capture data on the “what” and “how” program inputs were converted to outcomes and under what conditions throughout the implementation period. This paper presents the documentation of program processes (DPP) approach that is expected to help with understanding of how interventions were implemented to produce desired results. The (DPP) approach was developed to systematically capture how the program works, key contextual factors, and other program adaptations to produce results on an ongoing basis that could inform evaluations. The aim of DPP is to help program decision makers understand the “nuts and bolts” of what aspects of the program worked well and how they worked and avoid pitfalls that have undermined use of results from previous program evaluations.

The DPP includes tools and guidelines for programs to: 1) define the intervention package; 2) specify the theory of change; 3) plan and collect data on actual implementation and contextual factors that influence or hinder program processes, and any adaptations made; 4) develop a plan to regularly collect, analyze, and use data to address challenges and enhance program scale up.

**Objectives Sought:** The participants will understand that the DPP Approach was developed to:

1. Support efforts of scaling up and transferring successful programs to different settings;
2. Facilitate real time program learning, which allows for identification of bottlenecks and tracking of program adaptations and unintended consequences;
3. Provide the information necessary to interpret outcome results and make recommendations for program improvement.

**Brief Narrative and Justification:** The DPP approach is often talked about but rarely integrated and implemented to capture program processes in an ongoing basis on how the program works. The DPP will help design specific and well thought out evaluation questions during the design of the program, tracked on an ongoing basis, and helps set up the program for a rigorous evaluation that will contribute to the evidence base and are useful to decision makers. Good evaluation questions will fill a major gap in the program evaluation field by providing insights on how interventions work, the key influencing and hindering factors, any adaptation made to specific contexts, thereby contributing to the evidence base to help pave the way for program scale up and achieving development goals.

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**S 141 Documenting Program Process and Implementation****O 411 - A Stakeholder Centric Approach for Measurement Learning and Evaluation in Complex Interventions: The Complexity Responsive Stakeholder Centric Evaluation Framework***D. Chandurkar*<sup>1</sup><sup>1</sup> Sambodhi Research & Communications, Vice-president, Noida, India

By their very nature, complex system evaluations get embedded in a multi-stakeholder environment. From commissioners and implementers-the primary stakeholders, to policy-makers and global scientific and evaluation community-look for evidence, accountability and generalizable insights. Creating value for these stakeholders thus is onus of any such evaluation. The expectation is to meet the expectations at best. Given the boundary conditions that an evaluation works with-this is very challenging. However, this puts stakeholder expectations at the center of the design and implementation of such measurement, learning and evaluation (MLE) imperatives. And, necessitates plurality of approaches and integration of multiple methods so as to be comprehensive and responsive to needs of the multiple stakeholders.

Also, there are interventions that span very large geographies and multiple interventions are implemented at multiple levels-from grassroots to systems level. And stakeholder expectations range from impact measurement, learning about the context, pointers for scale-up and understanding intervention intersectionality. Need for plurality of approaches and methods gets underscored.

The presentation describes a **Complexity Responsive Stakeholder Centric Evaluation (CReSCE)** framework-for responding to the asks from a complex system evaluation keeping stakeholders at the loci. The framework maps the needs of the key stakeholders and builds on the MLE questions that needs to be responded from a stakeholder perspective. The interrelated components of the framework-**summative evaluation**, **process evaluation** and **complexity learning**-are envisaged to comprehensively address stakeholder needs.

The first component is of comprehensive **Summative Evaluation** to measure the net increments in key outcomes, and describe cost-effectiveness, sustainability and scalability of these effects. This responds to accountability concerns of the investors/commissioners and implementers and requires measurement of average treatment effect over the business as usual scenario. Sustainability here is also the concern of the investors while cost-effectiveness and scalability would be questions both from the investor as well as policy-makers.

Prerequisite to summative evaluation is assessment of quality of intervention and outreach; and understanding of the context. And, the second component of **Process Evaluation**. These periodic process studies would provide much necessary learnings for implementers so that mid-course corrections are facilitated.

The third component **Complexity Learning** explicitly focusses on learning and developing generalizable insights. The component is explanatory and exploratory-for understanding intervention-outcome linkages, outcome differentials in different contexts; and heterogeneity of implementation. And given the learning needs, requires experimentation with multiple approaches. An approach like Realist Synthesis can be experimented for understanding outcomes in different implementation contexts. Implementation Science frameworks can help understand determinants of implementation and implementation variability. Need-based studies may be required to learn about specific aspects of interventions issues, intersectionality in intervention sets, assumptions and causal linkages. Method and measure learning would also be a learning focus with evaluation and the scientific community as key stakeholders.

The proposed framework integrates the different components described above in comprehensively responding to MLE questions emanating from stakeholder needs. The **CReSCE** Framework therefore gets underscored as a responsive overarching framework that may be employed to guide design and implementation of the MLE imperatives.

Friday, 5 October 2018  
13:45–15:15

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 142 Evaluating Beyond The SDGs

## O 413 - Evaluating SDGs in No One Left Behind Lens

*A. Kalugampitiya*<sup>1</sup>

<sup>1</sup> EvalPartners, EvalPartners, Colombo 5, Sri Lanka

The paper is for participants to understand importance of “evaluating Sustainable Development Goals (SDGs) in no one left behind lens”. The workshop will bring the key learning perspectives of different stakeholders including evaluation professionals, public sector, private sector and the parliamentarians. The workshop is based on the guidelines: *Evaluating the Sustainable Development Goals With a “No one left behind” lens through equity-focused and gender-responsive evaluations* published by UN Women, EvalPartners in collaboration with other partners.

The 2030 Agenda for Sustainable Development comes together with a follow-up and review mechanism to ensure the SDGs are systematically monitored and reviewed to help countries implementing the 2030 Agenda to ensure “No one is left behind”. The follow-up and review mechanisms also call for inclusiveness, participation and ownership. This is why equity-focused and gender-responsive evaluation is needed. This transformative kind of evaluation can help countries to identify structural causes of inequalities through deeper analysis of power relationships, social norms and cultural beliefs. Integrating equity-focused and gender-responsive evaluations will provide strong evidence to ensure national voluntary reviews of SDGs are leaving no one behind.

Over the past few years, as civil society has become more active in development evaluation, a number of VOPEs have begun to have a formal consultative role in NEPSs and often provided the technical expertise. They also play an important role in generating demand for evaluations. VOPEs' role expands to advocate for equity focused and gender responsive evaluation as well.

National governments are the key agencies responsible for the implementation of the evaluations of SDGs within each country. As the reporting systems and evaluations are voluntary, the commitment of governments will be critical, particularly as they have to decide how to prioritize their limited financial and technical resources among many different development priorities – all supported by different groups of international and national stakeholders. Given the broad scope of the SDGs, almost all government agencies will potentially be involved and the national government will play an important coordinating role.

It is important to ensure that the evaluation strategies and approaches are fully consistent with the SDG principles for follow-up and review. This will ensure that the dimensions of gender and reducing inequalities are incorporated as an integral component of all of the SDG evaluations, and are not considered as special, stand-alone topics that are only of interest to gender specialists.

Gender equality, reducing inequalities and ensuring “No one left behind” are considered as distinct but linked core principles of the SDGs. Evaluations of SDGs policies should incorporate these principles throughout the evaluation in order to help address multiple causes of discrimination and exclusion. As a result of the work of gender equality advocates, gender equality is reflected throughout the 2030 Agenda, including in the declaration; goals, targets and indicators; means of implementation; global partnership; and follow-up and review.

Friday, 5 October 2018  
13:45–15:15

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 142 Evaluating Beyond The SDGs

## O 414 - The Importance of the 'S' in SDG: Lessons from a WASH Project in Ghana and Kenya

*C. Smit<sup>1</sup>, M. Jakoet<sup>1</sup>*

<sup>1</sup> Genesis Analytics, Evaluation for Development, Johannesburg, South Africa

Poor hygiene and sanitation, combined with water access from a source that is not protected from outside contamination (for example a stream or lake), can lead to the spread of water-borne diseases such as cholera, typhoid and diarrhoea, which in turn can contribute to other health issues such as malnutrition. This dual issue is thus a key risk factor contributing to morbidity and mortality in many countries around the world, and particularly in water-scarce African countries.

In response to this challenge, Sustainable Development Goal (SDG) 6, "Clean water and sanitation" aims to ensure access to safe water sources and sanitation for all, with the ultimate objective of promoting the intellectual capacity and productivity in a population through improved health and well-being. This paper will discuss the findings from an implementation evaluation of a water, sanitation and hygiene (WASH) project in Ghana and Kenya that aims to contribute to the achievement of SDG 6 in these respective countries. The paper will focus in particular on the importance of implementing realistic sustainability mechanisms early on to ensure the long-term sustainability of the project outputs and outcomes.

The purpose of the evaluation was to assess whether the project's operational mechanisms are likely to support the sustainable achievement of its objectives and, if not, to inform course correction to ensure that these are achieved during the remainder of the project period. Genesis used the OECD DAC criteria as the guiding framework for the evaluation, focusing specifically on Relevance, Effectiveness, Efficiency and Sustainability; and conducted site visits to each of the countries to collect data through key informant interviews and focus group discussions.

The projects in Ghana and Kenya are implemented by four implementing partners – two in each country – all of which follow slightly different approaches to addressing the challenge for their respective 'sub-project'. The evaluation found that each of the sub-projects is highly relevant and aligned to the needs of the target populations as well as the respective national priorities, which align to SDG 6. Additionally, the evaluation found that sub-projects are largely being implemented effectively and efficiently and are on track to achieve their objectives by the end of the grant period.

The key issue encountered by the evaluation is the potential sustainability of the projects beyond the period of support. WASH programmes have been implemented by governments and donors for many years, but ensuring long-term sustainability of water and sanitation infrastructure remains a challenge. The risk is that communities lack the capacity or support to maintain the water systems in question and in the event of a breakdown will revert to using previous, contaminated water sources. This session will discuss the importance of implementing partners having sustainability mechanisms in place but that the actual sustainability of a project hinges on critical assumptions that may not in fact hold in reality.

Friday, 5 October 2018  
13:45–15:15

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 142 Evaluating Beyond The SDGs

## O 415 - Citizen Participation and Inclusive Governance for Sustainable Development in Uganda: Lessons from Parliamentary Outreach Programs

*J. Watera*<sup>1</sup>

<sup>1</sup> Parliament of Uganda, Corporate Planning and Strategy, Kampala, Uganda

**Rationale:** The Sustainable Development Goals (SDGs) provide an opportunity for societies to engage in inclusive dialogues about the future together on the planet. The successful implementation of Agenda 2030 requires an integrated approach, partnerships and participation through inclusive political processes and responsive, effective, accountable institutions.

Parliamentarians have an opportunity, and a constitutional responsibility, to play a significant role in supporting and monitoring SDG implementation. The Agenda 2030 Declaration acknowledges the “essential role of national parliaments through their enactment of legislation and adoption of budgets, and their role in ensuring accountability for the effective implementation of SDG commitments.”

**Objectives:** Parliaments are uniquely positioned to act as an interface between the people and state institutions, and to promote and adopt people-centered policies and legislation to ensure that no one is left behind.

This paper intends to share the experience of outreach programs and to highlight some of the key achievements that the audience can draw lessons.

**Narrative and Justification:** Outreach programs are activities or programs involving engagement of stakeholders in an issue with the view of collecting their opinions. In the case of Uganda Parliament, it's the involvement of the citizenry at the grassroots levels in shaping the growth and development of the Country.

Parliament of Uganda has put in place a number of outreach programs, these include holding sessions of parliament at regional level, national schools debate hosted by Parliament, oversight field visits by committees of Parliament, Parliament week where Parliament opens up to the public, research week in Parliament, outreach programs by the Opposition office to assess the state of service delivery, schools visits, tournaments between sports teams in Parliament and the public, diaspora desk which takes stock of government responsiveness to the needs of Uganda in the diaspora among others.

Arising from the different avenues of parliamentary outreach programs, this paper will share the processes, the success stories, the challenges and lessons learned. This conference provides an opportunity upon which to explore citizen participation and inclusive governance approaches towards sustainable development in Uganda and therefore invites the audience to explore the take home points.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

## S 022 Turbulent Contexts and Evaluation

### O 416 - Sustainability Report of Municipality of Agios Dimitrios, Greece The society evaluates the local authorities

*I. Savvakis*<sup>1</sup>

<sup>1</sup> Municipality of Agios Dimitrios, Secretary General, Agios Dimitrios, Greece

Sustainability reporting enhances organizations' accountability for their impacts and therefore enhances trust, facilitating the sharing of values on which to build a more cohesive society.

With 2016 being the starting year for a sustainable development strategy, the Municipality of Agios Dimitrios issued its first Sustainability Report on November 2017. The composition of the Sustainable Development Report, according to international standards, makes the Municipality of Agios Dimitrios, the **first city in Greece**, but at the same time a pioneer on a pan-European level, as few cities have taken a similar initiative.

The beginning was through the adoption of the **10 principles of the United Nations Global Compact**, under which the City of Agios Dimitrios has pledged to work against corruption, with respect for human rights, work and the environment.

At the same time, the City has strategically decided to harmonize its actions, projects and initiatives with the 17 UN **Sustainable Development Goals (SDGs)** that seek to end poverty, protect the planet, and secure the prosperity of the whole of the world.

Subsequently, the Municipality **addressed all stakeholders** (such as citizens, municipality employees, partners, suppliers, agencies, associations, etc.) in order to identify and co-decide the material issues that need priority, following the parameters of the guidelines of the **GRI (Global Reporting Initiative)** and remaining faithful to the principles of objectivity and participation. Emphasizing the substantive issues has led to the development of the five pillars of the City's strategy for sustainable development.

- 1. Good Governance and Economic Sustainability,**
- 2. Environment,**
- 3. Society,**
- 4. Sustainable Infrastructures and**
- 5. Employees**

are the framework for activating, recording performance and reporting on the sustainable development of the Municipality of Agios Dimitrios, a continuous assessment of the administration of the municipality by the citizens.

Most importantly, these pillars lay the foundations for future sustainable strategies, taking into account the most vital aspects of everyday life, thus contributing to a better and more human city for all.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### S 047 Evaluation Capacity Building 1

## O 417 - How a VOPE comes true: RedEvalCR

A. Bolaños<sup>1</sup>, N. Salas<sup>1</sup>

<sup>1</sup> RedEvalCR, Central Commission, San José, Costa Rica

The creation of the Evaluation and Monitoring Network of Costa Rica (RedEvalCR) arises and is gestated due to the inexistence in the country of a democratic and articulated evaluation group, in which people with training, practice and interest in issues of evaluation felt represented. That lack of representation reduces the possibilities to growth as a guild, locally with a lack of knowledge of our needs and possibilities for development and integration, and externally, in a fragmentation and inadequate projection in different areas and sectors of development.

Faced with this reality, a group of people interested in the evaluation, raised the urgency of building a democratic, inclusive and open space that promotes dialogue and a horizontal collective action. This led to the start of a series of open calls in May 2017 for all people who felt called to the subject of evaluation, in order to create a space that would bring together the diversity of interests and knowledge of the people involved in the practice, research, professionalization and dissemination of the discipline of evaluation in Costa Rica.

As a result of a series of participatory workshops and complementary collaborative activities, RedEvalCR was born, as a non-profit organization, non-hierarchical, voluntary and joint work. This Network seeks to support and technically promote, strengthen, disseminate and expand the practice of a national evaluation culture, with the purpose of contribute with the social progress. Now on, RedEvalCR becomes a protagonist and formal technical reference of the practice and institutionalization of a Costa Rican evaluative culture.

This presentation exposes how the RedEvalCR understands the state of the art of the national evaluative task, and how from a collective and democratic construction, it is possible to generate an organizational structure of work, split into commissions, with responsibilities and specific tasks, whose results converged in a second stage, the establishment of an organizational structure and the final management model of RedEvalCR, as a formal non-profit organization, which provides an added value to the development of the Costa Rican evaluation culture.

# Panels

Wednesday, 3 October 2018  
10:00–11:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 01 The Role of Evaluation in the Face of Forced Displacement and the Emerging Humanitarian / Development Nexus**

## The Role of Evaluation in the Face of Forced Displacement and the Emerging Humanitarian / Development Nexus

M. Branco<sup>1</sup>, A. Tyrrell<sup>2</sup>

<sup>1</sup> [www.mariana-branco.com](http://www.mariana-branco.com), EES Board- EES Emerging Evaluators- EES Social Media, Porto, Portugal

<sup>2</sup> The World Bank, Independent Evaluation Group IEG, Dublin, Ireland

<sup>3</sup> UNHCR, Protection Office, Hatay, Turkey

**Rationale:** The issue of refugee response and integration is increasingly complex, highlighting the need for a comprehensive response. Moreover, under the umbrella of 2030 Agenda for Sustainable Development and following its principles and goals, in 2016 all Member States of the United Nations approved the New York Declaration for Refugees and Migrants. This panel will discuss how the evaluation community is handling the current humanitarian crisis and how evaluation can better embrace this call to reform.

**Justification:** “We are witnessing in today’s world an unprecedented level of human mobility. [...] there are roughly 65 million forcibly displaced persons, including over 21 million refugees, 3 million asylum seekers and over 40 million internally displaced persons.” (Article 3 of New York Declaration for Refugees and Migrants). The ever-growing number of forcibly displaced, the diminishing financial resources made available, the increasing number of violent conflicts since 2010 and the protracted nature of displacement pose a number of issues to the wider international community, all of which pose challenges to the evaluation community: (1) the agencies responsible for protecting displaced populations have been forced to broadening their mandate, dealing not only with asylum seekers, refugees and internally displaced persons (IDPs), but also all those in need of international protection, such as stateless people or migrants returning back to their countries of origin (people of concern). (2) the Declaration also clearly highlights the connection between humanitarian response and development, urging the creation of stronger cooperation between relevant actors that requires a comprehensive vision of the continuum between humanitarian action and development to address the needs of displaced persons and host States and communities. (3) Article 69 of this Declaration states the need for setting up a Comprehensive Refugee Response Framework (CRRF) on the basis of a “multi-stakeholder approach” that includes national and local authorities, international organizations, international financial institutions, civil society partners, the private sector, the media and refugees themselves.” (4) The adoption in 2018 of a Global Compact on Refugees demands for rethinking the way support is given to refugees, but also to local communities. It is crucial to ensure refugees have better access to health, education and livelihood opportunities and are included in their host communities from the very beginning. (5) Article 9 of the Zero draft version of the Global Compact on Refugees underlines that “...The programme of action is based on the recognition that humanitarian, development, and peace efforts are complementary and reinforce each other...”.

**Objectives:** Discuss: – The role of evaluation in face of a broadened mandate that encompasses a large spectrum of people of concern; – How can evaluation contribute to ongoing strategic reflections on the issue of humanitarian-development cooperation; – How can evaluation enhance the collaboration among development-oriented entities, international financial institutions, regional development – banks, private sector, impact investors, bilateral donors and non-traditional partners; – What is the role of evaluation in rethinking the way support is given to local communities; – How can evaluation enhance the complementarity between humanitarian aid, peacebuilding and peacekeeping.

**PN 01 The Role of Evaluation in the Face of Forced Displacement and the Emerging Humanitarian / Development Nexus****PC 001 - The Importance of the Humanitarian/Development Nexus in Realizing the SDGs – Challenges for Evaluation**

A. Tyrrell<sup>1</sup>

<sup>1</sup> The World Bank, Independent Evaluation Group IEG, Dublin, Ireland

Tony is a freelance evaluation consultant currently working with the Independent Evaluation Group (IEG) at the World Bank on a major evaluation on Forced Displacement. For the past six years Tony has worked with IEG on corporate level evaluations, country evaluations (e.g., Brazil, Tunisia) as well as various thematic, technical and learning pieces (e.g., Third Level Education, Service Delivery, Joint WBG Projects). Prior to this Tony worked with a number of consultancy companies on strategy and evaluation assignments, and with the ESF Evaluation Unit in Ireland where he was responsible for major evaluation on, for example, Early School Leaving, Local Development. Tony holds MA (UCD), and MSc (Trinity College, Dublin).

**PN 01 The Role of Evaluation in the Face of Forced Displacement and the Emerging Humanitarian / Development Nexus****PC 002 - Impact Investing and Humanitarian Aid: Can Value for Money Approaches Help Addressing the Gap?**

M. Branco<sup>1</sup>

<sup>1</sup> [www.mariana-branco.com](http://www.mariana-branco.com), EES Board- EES Emerging Evaluators- EES Social Media, Porto, Portugal

In 2016, the volume of international humanitarian assistance increased for the fourth year running, reaching a total of US\$27.3 billion. However: 1) while this amount is twelve times greater than fifteen years ago, never before has generosity been so insufficient – currently, there is a funding gap for humanitarian action of an estimated US\$ 15 billion; 2) the evidence base for designing humanitarian programs is sparse and the knowledge vacuum limits our capacity to draw effective interventions, making of rigorous evaluation a matter of paramount importance; 3) ensuring enough quality money for humanitarian crises is not just about writing a bigger cheque – it's about increasing the cost-benefit and value for money of the operations; 4) beyond focusing assistance on fragile countries, there has to be systematic investment in resilience-building, which includes dedicated funds for peacebuilding and conflict resolution. On the other hand, the development of impact investing and social finance at a global scale signal a growing trend of importing methods and logics of the business sector to the third sector, while blending financial and social values. Initiatives such as The Impact Investing Forum for INGOs, The Impact Evaluation to Development Impact (i2i) and Peace Nexus are examples of how these tendencies are reaching out to the humanitarian sector. Nevertheless, challenges arise when combining the two worlds. Conflict prone contexts and rapid changing settings demand for adjusted methods; tradition to focus on outputs undermines the transition to outcomes and impact management; fatigue and sense of guilt generate aversion to failure; need for cooperation instead of competition; etc. make it difficult, for financial and humanitarian actors, to communicate in the same language. Traditional literature suggests that impact evaluation and cost-benefit analysis can help suppressing these gaps and more recent studies report a consensus amongst major donors and aid agencies that considering Value for Money (VfM) approaches, is worth trying. However more reflection is needed regarding the operationalization of VfM in the context of humanitarian sector. This presentation discusses how VfM methods (frameworks that help to account and maximizing the impact of each monetary unit spent to improve people's life) can help addressing the humanitarian financing gap, while providing practical evidence that helps managers optimize the creation of social value towards the achievement of the SDGs. This paper consists of an exploratory study that triangulates literature review, experience applying VfM approached and field experience evaluating the humanitarian sector. Open questions and handouts will be distributed in order to stimulate a participative audience.

Brandstetter and Lehner. 2015. Opening the Market for Impact Investments: The Need for Adapted Portfolio Tools. *Entrepreneurship Research Journal*. Christoplos. 2016. The evolution of humanitarian Evaluation. *Global Humanitarian Assistance*. 2017. *Global humanitarian assistance report 2017*. The World Bank. 2017. *i2i Annual Report 2017*. World Humanitarian Summit. 2016. High-Level Panel on Humanitarian Financing Report to the Secretary-General: Too important to fail – addressing the humanitarian financing gap.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

PN 02 Stories and Scars From the Field: Guidelines for Managing Conflict in the Conduct of Evaluation

## Stories and Scars From the Field: Guidelines for Managing Conflict in the Conduct of Evaluation

*J. Owen<sup>1</sup>, S. Bayley<sup>2</sup>, R. Cummings<sup>3</sup>, B. Perrin<sup>4</sup>*

<sup>1</sup> The University of Melbourne, Centre for Program Evaluation, Melbourne, Australia

<sup>2</sup> Australian Government, Department of Foreign Affairs and Trade, Canberra, Australia

<sup>3</sup> Murdoch University- Australia, Emeritus Professor, Perth, Australia

<sup>4</sup> Consultant, Consultant, La Masque, France

Objectives. The aim of this panel are to: \*examine ways in which conflict affects the conduct of evaluation practice, and to \*suggest methods by which evaluators can better handle and resolve conflicting situations. Narrative. There is broad agreement about the political nature of evaluation. Theorists such as Carol Weiss[1] note that evaluation is conducted within a context of competing pressures of interests, ideologies, and institutional constraints. Such pressures can lead to conflicting situations that an evaluator must be equipped to handle. Conflict in evaluative work falls within the domain of people skills. Where and how do practising evaluators obtain skills to handle these situations? There are claims that the evaluation profession has not supported the inclusion of these skills in formal and informal training programs. This panel is designed to contribute to such a body of knowledge by examining recent good practice. Three experienced presenters (from Australia) will each provide a case that has arisen in their recent work, and each will outline how conflict was resolved. In addition, tentative common threads across the cases will be introduced. Attendees will be encouraged to add examples from their own experiences. A respondent (with experience in European settings) will comment on the core presentations and the salience of common threads across the cases. We expect that this will lead to analysis of the nature of conflict that allows for variation in approach at different stages of the evaluation process; during the (i) negotiation/planning (ii) data management, and (iii) findings/reporting stages. Rationale. We are aware that the social science literature contains extensive literature both on conflict and conflict resolution between social groups and in managing change. Through the presentations in this panel, we aim to extend the limited emerging conceptual understandings about conflict that apply to the field of policy and program evaluation, for example those by King and Stevahn [2]. We expect that panel deliberations will contribute to a body of knowledge that is particularly applicable to the effectiveness of future of evaluation work internationally. Review Criteria. Our view is that this panel touches on all of the abstract review criteria. However, in particular the panel will make specific contributions to two criteria in particular, in ways that other conference presentations may not. These are Creativity and Innovation and Resilient Societies. \*Creativity and Innovation. Conflict with evaluation practice is an innovative issue and should be of interest to conference attendees. \*Resilient Societies. The ability of evaluators and the evaluation community to understand and cope with conflict situations improves the resilience of practitioners. Thus, this panel deals with knowledge and skills to enable evaluators to cope with an area of practice that is not usually covered in programs of evaluator preparation. [1] Weiss C. The Interface between evaluation and public policy evaluation. *Evaluation*, 1999 5: 4; 468–486. [2] King JA and Stevahn, L. (Managing conflict constructively in evaluation settings. *Interactive Evaluation Practice: Mastering the Interpersonal Dynamics of program evaluation*: 2014:166–194. Sage Publications

**PN 02 Stories and Scars From the Field: Guidelines for Managing Conflict in the Conduct of Evaluation**

**PC 004 - Conflict within Evaluation of the Improving Schools Program**

J. Owen<sup>1</sup>

<sup>1</sup> The University of Melbourne, Centre for Program Evaluation, Melbourne, Australia

Structure of Session (Owen); 10 minutes.

**PN 02 Stories and Scars From the Field: Guidelines for Managing Conflict in the Conduct of Evaluation**

**PC 005 - Conflict within Evaluation of the Managing for Development Results Strategy**

S. Bayley<sup>1</sup>

<sup>1</sup> Australian Government, Department of Foreign Affairs and Trade, Canberra, Australia

Case 1 (Bayley): Conflict within Evaluation of the Managing for Development Results Strategy (15 minutes).

**PN 02 Stories and Scars From the Field: Guidelines for Managing Conflict in the Conduct of Evaluation**

**PC 006 - Conflict within Evaluation of the Western Australian Tough on Graffiti Strategy**

R. Cummings<sup>1</sup>

<sup>1</sup> Murdoch University- Australia, Emeritus Professor, Perth, Australia

Case 2 (Cummings): Conflict within Evaluation of the Western Australian Tough on Graffiti Strategy (15 minutes).

**PN 02 Stories and Scars From the Field: Guidelines for Managing Conflict in the Conduct of Evaluation**

**PC 007 - Conflict within Evaluation of the Improving Schools Program**

J. Owen<sup>1</sup>

<sup>1</sup> The University of Melbourne, Centre for Program Evaluation, Melbourne, Australia

Case 3 (Owen): Conflict within Evaluation of the Improving Schools Program (15 minutes).

**PN 02 Stories and Scars From the Field: Guidelines for Managing Conflict in the Conduct of Evaluation**

**PC 008 - Reviewing Conflict Across the Case Studies and the Literature**

B. Perrin<sup>1</sup>

<sup>1</sup> Consultant, Consultant, La Masque, France

Review and Analysis of Contributions (15 minutes).

Wednesday, 3 October 2018  
10:00–11:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 03 There is no Resilience without Equity: When will our Profession Finally Act to Reverse Asymmetries in Global Evaluation?**

## **There is no Resilience without Equity: When will Our Profession Finally Act to Reverse Global South Asymmetries in Global Evaluation?**

A. Sibanda<sup>1,2</sup>

<sup>1</sup> AFREA, IOCE, President, Johannesburg, Zimbabwe

<sup>2</sup> S2SE, chair

Launched by the five Global South Evaluation Networks from Asia, Africa, Latin America, Asia-Pacific and the Caribbean, the **South to South Evaluation Initiative (S2SE)** seeks to address the fundamental inequalities and asymmetries in knowledge, evidence, resources and capacity between the Global South and the Global North. S2SE asserts that while good progress has been made in building the global evaluation architecture, evaluation theory and practice remains overwhelmingly dominated by the philosophies, worldviews, methods and priorities of the Global North.

This dynamic panel of evaluation leaders from the Global South challenges EES and the global evaluation system to recognize and address these asymmetries. Global South evaluation leaders will provide examples of how EES commissioners of evaluation, practitioners, policy makers, funders and academics can ensure that all parts of the world contribute substantively to evaluation and global development.

**Please come and be part of this long overdue Call to Action!**

**PN 03 There is no Resilience without Equity: When will our Profession Finally Act to Reverse Asymmetries in Global Evaluation?**

## **PC 009 - Hidden Power and Privilege in Evaluation: A Latin American Perspective**

S. Salinas<sup>1</sup>

<sup>1</sup> Latin American and Caribbean Monitoring, Evaluation and Systematization Network (ReLAC), Coordinator, La Paz, Bolivia

Silvia will highlight how Latin American evaluation leaders are using post-colonial and decolonization theory to reconsider and redesign evaluation approaches to explicitly address the hidden power and privilege that is embedded in evaluation and development.

**PN 03 There is no Resilience without Equity: When will our Profession Finally Act to Reverse Asymmetries in Global Evaluation?**

## **PC 010 - South Asian Evaluators Tackle Power Asymmetries in South Asia**

S. Zaveri<sup>3</sup>

<sup>1</sup> South Asia Community of Evaluators (CoE), Vice Chair, Mumbai, India

Sonal will highlight how South Asian evaluators are taking steps to decolonize and indigenize evaluation thinking and practice in South Asia. She will discuss how the realities and challenges of 'Push and Pull' factors in the increasingly globalized context of South Asia pay little attention to local realities that perpetuate historical inequities and strengthen concentrations of power.

**PN 03 There is no Resilience without Equity: When will our Profession Finally Act to Reverse Asymmetries in Global Evaluation?**

## **PC 011 - Made in Africa Evaluation: Why Africa-Centric Evaluation Matters in Global Evaluation**

Z. Ofir<sup>1</sup>

<sup>1</sup> IDEAS and IOCE, Vice-President, Johannesburg, South Africa

<sup>2</sup> AEA, Board member

<sup>3</sup> AfrEA, former President

Zenda will trace the evolution of Africa-rooted and Africa-led M&E, explaining the journey of African evaluation leaders and scholars in re-examining the paradigms that guide evaluation practice on the continent, the importance of African values and worldviews, and how 'Made in Africa Evaluation' is fostering intellectual leadership, knowledge, capacity and scholarship in Africa and beyond.

**PN 03 There is no Resilience without Equity: When will our Profession Finally Act to Reverse Asymmetries in Global Evaluation?**

## **PC 012 - Funders Can Do Better and They Know It: What Is Stopping Them?**

N. MacPherson<sup>1</sup>

<sup>1</sup> Independent, New York, USA

<sup>2</sup> Rockefeller Foundation, former Evaluation Director, New York, USA

Nancy will highlight ways in which funders and investors, individually and collectively, can and should play a critical role in reversing inequalities and asymmetries in knowledge, capacity, resources and the influence of the Global South in data, evidence and decisions affecting the Global South. She challenges the funding and investment community that they can and should do better.

**PN 03 There is no Resilience without Equity: When will our Profession Finally Act to Reverse Asymmetries in Global Evaluation?**

## **PC 285 - Who holds the Power in the Evaluation System? Perspectives from the Global South**

A. Sibanda<sup>1,2</sup>

<sup>1</sup> AFREA, IOCE, President, Johannesburg, Zimbabwe

<sup>2</sup> S2SE, chair

Adeline will discuss the urgency of the South to South Evaluation Initiative's Call to Action and highlight questions raised by Global South evaluators as they seek to address power asymmetries in the global evaluation system.

Wednesday, 3 October 2018  
10:00–11:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 04 Evaluation for the Anthropocene: A Call to Action for a Sustainability-Ready Evaluation**

## Evaluation for the Anthropocene: A Call to Action for a Sustainability-Ready Evaluation

A. Rowe<sup>1</sup>

<sup>1</sup> ARCEconomics, British Columbia, Director, Maple Bay, Canada

Incorporating sustainability into evaluation requires a significant shift in the ways we think about, undertake and commission evaluations. The aim of the is to help participants understand the difference a sustainability-ready evaluation will make to your work, to help you get started and to contribute to an emerging checklist for sustainability-ready evaluation. The session will take a working group and capacity building approach. We will first introduce the concept of sustainability-ready evaluation and the need for sustainability to be a cross-cutting issue, and then illustrate the consequences of ignoring natural systems in evaluations based in the human system and ignoring the human system in evaluations based in natural systems. Most of the session time will be devoted to working in smaller groups and as then as a whole to consider whether and how sustainability applies to evaluation examples nominated by participants and to understanding why sustainability might not have been considered in many of these examples. We will then together develop items for an initial checklist for sustainability-ready evaluation and identify the gaps in evaluation knowledge, methods and remit and the challenges sustainability-ready presents to the worldview of evaluation and the connectivity to Indigenous evaluation. The panel are thought and practice leaders in the emerging concerns of sustainability in evaluation and of evaluations of complex evaluations conducted in coupled human and natural system settings.

**PN 04 Evaluation for the Anthropocene: A Call to Action for a Sustainability-Ready Evaluation**

## PC 014 - Panel Contribution

A. Rowe<sup>1</sup>

<sup>1</sup> ARCEconomics, British Columbia, Director, Maple Bay, Canada

**PN 04 Evaluation for the Anthropocene: A Call to Action for a Sustainability-Ready Evaluation**

## PC 015 - Panel Contribution

J. Uitto<sup>1</sup>

<sup>1</sup> Global Environment Facility, Independent Evaluation Office, Washington- DC, USA

**PN 04 Evaluation for the Anthropocene: A Call to Action for a Sustainability-Ready Evaluation**

## PC 016 - Panel Contribution

A. Brousselle<sup>1</sup>

<sup>1</sup> University of Victoria, School of Public Administration, Victoria, Canada

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 05 Leveraging Longitudinal Data for Resilience Measurement**

## Leveraging Longitudinal Data for Resilience Measurement

*B. Sagara*<sup>1</sup>

<sup>1</sup> Mercy Corps, Research and Learning, Oakland, USA

Measuring resilience requires the ability to observe how households respond, cope, adapt and recover to shocks and stresses over time. Consequently, having longitudinal data is fundamental to measuring and evaluating resilience programs. While panel data has long been utilized across various disciplines, applying it to international development projects focused on building resilience is relatively new. The objectives of this panel are to: 1. Present three examples of how longitudinal studies are being designed, implemented, and analyzed to measure program contributions to resilience 2. Discuss promising practices, challenges, and new frontiers in panel survey designs for resilience measurement In addition to chairing the panel, Brad Sagara will share experiences and insights leveraging World Bank LSMS panel data from Nigeria to understand conflict and household resilience. Most research and evaluation to date focuses on resilience to natural disasters, with considerably less focused on resilience to the effects of conflict. Using three waves of nationally representative panel data spanning the rise of Boko Haram and ongoing conflict in the Delta region, this research aims to understand both the impact of conflict on key measures of household wellbeing and understand the correlates to improved wellbeing outcomes in the context of conflict.

**PN 05 Leveraging Longitudinal Data for Resilience Measurement**

## PC 017 - Insights and Lessons Learned From High Frequency Recurrent Monitoring in Ethiopia

*T. Frankenger*<sup>1</sup>

<sup>1</sup> TANGO International, Consulting Firm, Tucson, USA

Tim Frankenger will share the results and insights from a panel of households in two districts of Ethiopia, selected to observe how households were managing and recovering from multiple severe waves of drought. Using mixed-methods panel data, the research aimed to understand various dynamics of shock and shock exposure, what household responses were most effective, and how wellbeing, household, and community characteristics, changed over time. Likely the first of its kind, Tim will share key lessons learned from this study design and how they have been adapted in subsequent iterations across Sub-Saharan Africa and Southeast Asia.

**PN 05 Leveraging Longitudinal Data for Resilience Measurement****PC 018 - Two Years Later: Determinants of Coping and Recovery to the Gorkha Earthquake in Nepal**

*J. Scantlan<sup>1</sup>*

<sup>1</sup> *Mercy Corps, Southeast Asia Regional Resilience Hub, Kathmandu, Nepal*

Jill Scantlan will share the results and insights from a panel of households surveyed 10-weeks, 1-year, and 2-years after the 7.8 magnitude earthquake that struck Nepal in April 2015, killing over 9,000 people, destroying or badly damaging more than 800,000 homes, and displacing approximately 2.8 million people. The analysis explored what factors mattered most for short-term coping and long-term recovery to improve humanitarian response and design of recovery programs in the aftermath of acute disasters. In addition to this novel research, Jill will also share how she is working with the USAID-funded project Promoting Agriculture, Health, and Alternative Livelihoods (PAHAL) in Nepal to integrate qualitative methods to follow households as they experience and respond to shocks and stresses. This approach will document rich experiences of resilience in Nepal with the aim of understanding what builds resilience and determines resilience trajectories at the individual, household, community and system level.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 06 How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals and International NGO Evaluators****How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals and International NGO Evaluators***I. Davies<sup>1</sup>*<sup>1</sup> *Ian C Davies Conseil Inc, Chief Executive Officer, Victoria, Canada*

All international development organizations in one way or another aim to contribute to the 'sustainable development' of societies and citizens. This is most recently evidenced by 2015–2030's Sustainable Development Goals as well as the Impact Investor boom aimed at social impact. In this context, evaluators routinely ask questions around how well are our international development efforts (through our funding, design, implementation, monitoring and evaluation of projects, programmes and partnerships) achieving sustained results in the short/medium term that countries can built on? The panel of evaluation experts from different parts of the international development community discusses a number of critical questions on this seemingly routine yet inherently complex question. First, what do we mean by sustainability? Second, while long-term sustainability is seen as more the responsibility of the developing country government, what is the role of donors and implementers in fostering sustainability and resilience when we give aid? Finally, what are the evaluation modalities that we are using or should be using to evaluate sustainability to improve learning and accountability on this crucial issue? Specific contributions of panelists: Scott Bayley (Australia's DFAT): Evidence from ADB and the Australian Department of Foreign Affairs and Trade illustrating that by not undertaking post project evaluations of sustainability we miss the opportunity to learn what worked, what didn't and why; and hence we won't understand how to improve our interventions over time, and what policies are needed to strategically position the aid industry. Rob D. van den Berg (King's College London): Sustainability needs to be looked at from a country and portfolio systems perspective rather than as a continuation of benefits or outcomes from a project perspective. Useful evaluation modalities in this regard include ex-post portfolio evaluations, country programme evaluations, thematic evaluations and policy evaluations. Key is the diminishing status of development cooperation, the recognition that development is now a global issue, and that one of the greatest challenges we face is to move towards global sustainability. Jindra Cekan/ova (Valuing Voices): Presents evidence from ex-post (sustained and emerging impacts) evaluations 2+ years after closeout by INGOs, USAID, JICA and OECD. Few projects are evaluated, and those that are have very mixed results which indicate short-term vision and issues of low accountability of projects to country-nationals' successful uptake. Yet some evidence of what works and how to do more of it does exist and shows that getting to sustained impacts need much longer funding, collaborative design, adaptive implementation and many more lessons. Marco Segone (UN-FPA): "Real-world" perspective of development agencies, explaining what are the challenges and potential ideas for the way forward. Jos Vaessen (World Bank): What are some of the key conceptual and methodological ingredients of a sustainability perspective in evaluation? How can theory-based and systems perspectives (possibly connected to the former) contribute to a better understanding of sustainability issues. What are the implications for different evaluation modalities? Ian Davies (Ian Davies, Conseil, CHAIR): Results, e.g. impacts, cannot be sustained but processes may be sustained. And even they, and their sustainability, are dependent on context.

**PN 06 How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals and International NGO Evaluators**

**PC 019 - How Sustained and Resilient are our Impacts?  
Perspectives from Multilaterals, Bilaterals and International  
NGO Evaluators: Views from the Chair**

I. Davies<sup>1</sup>

<sup>1</sup> Ian C Davies Conseil Inc, Chief Executive Officer, Victoria, Canada

As Chair, Ian Davies will challenge panelists whether results, e.g. sustained impacts, can or cannot be sustained. He holds that processes may be sustained. And even they, and their sustainability, are dependent on context.

**PN 06 How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals and International NGO Evaluators**

**PC 020 - How Sustained and Resilient are our Impacts?  
Perspectives from Multilaterals, Bilaterals and International  
NGO Evaluators: Views from NGOs and Participants**

J. Cekanova<sup>1</sup>

<sup>1</sup> Valuing Voices at CEKAN CONSULTING LLC, President/ Catalyst, Prague, Czechia

Jindra Cekan, PhD presents evidence from ex-post (sustained and emerging impacts) evaluations 2+ years after closeout by INGOs, USAID, JICA and OECD. Few projects are evaluated, and those that are have very mixed results which indicate short-term vision and issues of low accountability of projects to country-nationals' successful uptake. Yet some evidence of what works and how to do more of it does exist and shows that getting to sustained impacts need much longer funding, collaborative design, adaptive implementation and many more lessons.

**PN 06 How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals and International NGO Evaluators**

**PC 021 - How Sustained and Resilient are our Impacts?  
Perspectives from Multilaterals, Bilaterals, International  
NGO Evaluators: A View from ADB & AusAid**

S. Bayley<sup>1</sup>

<sup>1</sup> Australian Department of Foreign Affairs and Trade, Performance Management and Results, Canberra, Australia

Scott Bayley presents evidence from ADB and the Australian Department of Foreign Affairs and Trade illustrating that by not undertaking post project evaluations of sustainability we miss the opportunity to learn what worked, what didn't and why; and hence we won't understand how to improve our interventions over time, and what policies are needed to strategically position the aid industry.

**PN 06 How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals and International NGO Evaluators**

**PC 022 - How Sustained and Resilient are our Impacts?  
Perspectives from Multilaterals, Bilaterals and International NGO  
Evaluators: the World Bank/IEG & Maastricht**

J. Vaessen<sup>1</sup>

<sup>1</sup> World Bank, Independent Evaluation Group, Washington- D.C., USA

Jos Vaessen, PhD of the World Bank/IEG Group and Maastricht University asks: What are some of the key conceptual and methodological ingredients of a sustainability perspective in evaluation? How can theory-based and systems perspectives (possibly connected to the former) contribute to a better understanding of sustainability issues. What are the implications for different evaluation modalities?

**PN 06 How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals and International NGO Evaluators**

**PC 023 - How Sustained and Resilient are our Impacts?  
Perspectives from Multilaterals, Bilaterals and International NGO  
Evaluators: A View from Development Agencies**

M. Segone<sup>1</sup>

<sup>1</sup> United Nations Population Fund – UNFPA, Director- Evaluation Office, New York, USA

Marco Segone of UNFPA and EvalPartners, among others, reminds us of “real-world” perspective of development agencies, explaining what are the challenges and potential ideas for the way forward toward sustained impact and sustainable development.

**PN 06 How Sustained and Resilient are our Impacts? Perspectives from Multilaterals, Bilaterals and International NGO Evaluators**

**PC 024 - How Sustained and Resilient are our Impacts?  
Perspectives from Multilaterals, Bilaterals and International  
NGO Evaluators: A Bilateral and Academic View**

R.D. van den Berg<sup>1</sup>

<sup>1</sup> King's College London, Department for International Development, London, United Kingdom

Rob D van denBerg provides a broad overview that sustainability needs to be looked at from a country and portfolio systems perspective rather than as a continuation of benefits or outcomes from a project perspective. Useful evaluation modalities in this regard include ex post portfolio evaluations, country programme evaluations, thematic evaluations and policy evaluations, he will discuss how we can move towards sustained results, and what challenges we must overcome.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 07 Tackling Complexity Theory in Evaluation: Methodological Challenges, Lessons Learned and Practical Solution**

## **Tackling Complexity Theory in Evaluation: Methodological Challenges, Lessons Learned and Practical Solutions**

*C. Maldonado Trujillo*<sup>1</sup>

<sup>1</sup> *Center for Research and Teaching in Economics- Center for Learning on Evaluation and Results Latin America, Public Administration Division, Mexico city, Mexico*

Complexity has long been a part of the evaluative profession and discourse. However, complexity theory, in its various strands, has only recently been explicitly incorporated into public policy discourse, evaluation theory and interrelated methodological debates. This panel offers a thorough conceptual review of the theoretical implications of complexity theory for evaluation methodology, practice and ethos as a starting point to inform practical strategies to cope with the formidable challenges entailed by recent developments in the field, and the explicit adoption of the complexity and resiliency lens in development interventions ( e.g. SDGs). The panel is composed by a diverse group of international scholars and practitioners that jointly present an internationally-inspired and policy-diverse theoretical perspective with a common practical concern. The panel is composed by Claudia Maldonado (Mexico, Chair), María Bustelo (Spain), Michael Bamberger (USA) and Tessie Catsambas (USA/Greece) and will have Thomas Schwandt, a renowned and experienced evaluation theorist and practitioner as discussant. Each participant will have 15 minutes for his/her presentation. The starting point of this panel is that theoretical clarification and a common conceptual ground are needed to effectively incorporate complexity in evaluation theory and practice. However, while recognizing the great challenge that the adoption of a complexity perspective represents, we jointly argue that there are practical and viable ways for the evaluation community to confront it.

The first presentation, by Claudia Maldonado, focuses on theoretical and conceptual clarification: it offers an account of the origins, scope and limitations of complexity theory and some of its implications for evaluation. The discussion focuses on how the theories of complexity question classical notions of causality, empirical identification strategies and highlights the need to relax and review key assumptions of evaluation theory, and practice. Secondly, María Bustelo showcases the rich trial and error experience of gender mainstreaming and structural change both from an academic and decision-maker perspective, drawing on the many lessons that can and should be incorporated in the policy dialogue around complexity. She analyzes the challenge of evaluating resistance to change within institutions driving that change, inconsistencies that are encountered while trying to transform power relations and gender hierarchies, as well as the complexity of joint action and the necessary inclusion of diverse voices, agencies and actors. Finally, two complementary, and practically-oriented visions are presented. Michael Bamberger draws on his extensive experience as a methodologist and practitioner in order to present a practical proposal of five steps to make complexity-sensitive evaluation with limited resources, and Tessie Catsambas presents a practitioner's perspective on the road towards SDGs-responsive evaluation that focuses on the way it enables evaluation stakeholders to participate and interact in a substantive way with evaluation. Jointly, these participations offer diverse perspectives on complexity that balance theoretical nuance with practical roadmaps, drawing on Scriven's notion that in evaluation there is nothing more practical than a good theory.

**PN 07 Tackling Complexity Theory in Evaluation: Methodological Challenges, Lessons Learned and Practical Solution****PC 025 - Tackling Complexity: Gender Mainstreaming as a Roadmap**

M. Bustelo<sup>1</sup>

<sup>1</sup> Complutense University of Madrid, Political Science and Public Administration, Madrid, Spain

This individual paper argues that many of the theoretical implications of complexity have already been part of the conceptual and methodological baggage of gender mainstreaming policies and gender-sensitive evaluation. Accordingly, it draws on the lessons learned from gender mainstreaming as a showcase of complex interventions aimed at promoting structural change. This contribution draws both from extensive academic evidence on gender mainstreaming, focusing on the specific challenges for evaluation, as well as a decision-makers perspective from the author as a driver of gender equality at a large institution. It presents the challenges confronted by gender mainstreaming and some of their implications for complexity-sensitive evaluation.

**PN 07 Tackling Complexity Theory in Evaluation: Methodological Challenges, Lessons Learned and Practical Solution****PC 026 - Complexity-Sensitive Evaluation: A Five Step Approach**

M. Bamberger<sup>1</sup>

<sup>1</sup> Independent International Consultant, Evaluation and Social Policy, Beaverton, USA

The starting premise of this paper is that acknowledging the need and the implications of complexity theory for evaluation should not lead to the potential paralysis of evaluation efforts. To the contrary, instead of demanding heroic assumptions on evaluation capacity and resources or overestimating the methodological challenges posed by complexity theory, this contribution presents a five step practical approach to tackle complexity in real world scenarios. It argues that a creative unpacking of complex interventions that borrows from various well-known and tested methodologies is an actionable methodological solution to a pressing theoretical challenge faced by the evaluation community.

**PN 07 Tackling Complexity Theory in Evaluation: Methodological Challenges, Lessons Learned and Practical Solution****PC 027 - SDG-Responsive Evaluation: Are We Getting Better at Managing Complexity?**

A. Tzavaras Catsambas<sup>1</sup>

<sup>1</sup> EnCompass, Evaluation, Rockville, USA

The adoption of the SDGs invited the evaluation community to think deeply about evaluating in an SDG-responsive way. Two core elements of SDG-responsive evaluation are (1) the way evaluation embraces and addresses complexity, and (2) the way it enables evaluation stakeholders to participate and interact in a substantive way with an evaluation. Michael Bamberger's book *Dealing with Complexity in Development Evaluation* addresses the first element well presenting strategies on how to address complexity in designing evaluations and addressing methodology challenges in evaluating social programs. The second element is only partly addressed in approaches such as transformative evaluation, appreciative evaluation, gender-responsive evaluation. The question is: how should the evaluator's role change, to ensure that evaluation should leave no one behind? In what ways must the evaluator insist in inviting evaluation participants who would normally not be part of the evaluation process?

**PN 07 Tackling Complexity Theory in Evaluation: Methodological Challenges, Lessons Learned and Practical Solution****PC 028 - Striking a Balance Between Theory, Methodological Concerns and Practical Solutions for Evaluation?**

*T. Schwandt*<sup>1</sup>

<sup>1</sup> University of Illinois, Department of Education, Urbana, USA

As a discussant, this contribution focuses on the critical review of the individual panelists' participation and the overall contribution of their complementary perspectives to the proposed aim of this panel, namely tackling complexity with a theoretically-informed, yet practically oriented approach to evaluation.

**PN 07 Tackling Complexity Theory in Evaluation: Methodological Challenges, Lessons Learned and Practical Solution****PC 029 - Implications of Complexity Theory for Evaluation**

*C. Maldonado Trujillo*<sup>1</sup>

<sup>1</sup> Center for Research and Teaching in Economics- Center for Learning on Evaluation and Results Latin America, Public Administration Division, Mexico city, Mexico

This paper focuses on theoretical and conceptual clarification: it offers an account of the origins, scope and limitations of complexity theory and some of its main implications for evaluation theory and practice. The discussion focuses on how the theories of complexity question classical notions of causality, how they affect empirical identification strategies and highlights the need to relax and review key assumptions of evaluation theory and practice. It argues that a common language and understanding of complexity theory is needed in evaluation and policy dialogue, and that methodological adjustments and professional practice need to be revised in order to effectively adopt a complexity perspective.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 08 How Can Evaluation Help Bureaucracies Be, Well, Less Bureaucratic?**

## **How Can Evaluation Help Bureaucracies be, Well, Less Bureaucratic?**

*B. Perrin<sup>1</sup>, T. Tyrrell<sup>2</sup>*

<sup>1</sup> *Independent consultant, independent, Vissec, France*

<sup>2</sup> *Independent consultant, Dublin, Ireland*

When democratic institutions come under threat in turbulent times, they are increasingly perceived as remote and disconnected from ordinary people. There are strong views that many bureaucratic requirements, based upon practices and compliance with rules and procedures, inhibit rather than facilitate effective public services and meaningful accountability. How can evaluation support appropriate bureaucratic practices? Or, are some evaluation approaches part of the problem rather than part of the solution?

This session will provide participants with an opportunity to actively engage in a discussion about the intent of bureaucracy, the purpose and utility of many bureaucratic practices that may be viewed as impediments rather than supports to responsive public services. The role of evaluation at its intersection with bureaucratic practices could be made more effective. The panel will take a critical look at evaluation practice in the context of bureaucracies, suggesting that some evaluation approaches and practices may be part of the problem rather than supporting bureaucratic practices to be well aligned to deliver more effective, impactful public services that are seen by citizens as responsive to their needs. This issue relates to the conference theme of how evaluation can better support resilience and democratic, although bureaucratic organisations currently under threat in turbulent times. For example, part of the disenchantment of some European citizens with the EU relates to perceptions of unnecessary bureaucracy, with even strong supporters of the EU, such as France's President Macron, saying that it is overly bureaucratic. This session will look through an evaluative lens at various aspects and practices of bureaucracy. How can evaluation contribute to improvements in bureaucratic practices, such as in taking an empirical look at bureaucratic processes and identifying how it could potentially lead to improvements. On the other hand, are some evaluation practices themselves merely bureaucratic rituals that make things worse? The panel will present examples of the interaction between evaluation and bureaucracy. As professional evaluators, we know that questions such as "Is bureaucracy good or bad?" are too simplistic. Rather, this session will explore questions such as: What can an evaluative lens tell us about varying characteristics of bureaucracy? Under what circumstances do various practices enhance democracy and benefits to citizens? Under what circumstances might they not be so? We may distinguish between bureaucracy as an organisational form or structure vs. bureaucratic practices that may or may not always serve their intended purpose. Rather than a typical panel of 'talking heads', this will be a highly participatory session. The panel members will give short presentations, raising issues for group discussions, each with a lead panellist as follows: Penny Hawkins (co-chair) – Why focus interest on bureaucracy and the roles of evaluation? Burt Perrin (co-chair) – The disconnect between outcome-oriented public services and activity-based bureaucratic practices. Jacques Toulemonde – Management paradigms, accountability, evaluation, and bureaucracy in (un)predictable contexts. Alison Pollard – Evaluation quality assessment: useful tools vs. bureaucratic impediments to learning)? Francesco Mazzeo Rinaldi – Disconnect between monitoring/evaluation indicators in the EU cohesion policy.

**PN 08 How Can Evaluation Help Bureaucracies Be, Well, Less Bureaucratic?****PC 030 - Why Interest in Bureaucracy and Potential Roles for Evaluation**

*T. Tyrrell<sup>1</sup>*

<sup>1</sup> *Independent consultant, Dublin, Ireland*

Penny will introduce the session, and the panelists. She will indicate why we are interested in bureaucracy, and the potential role that evaluation could play in assessing the appropriateness of bureaucratic practices. She will explain how the session will actively engage all participants in identifying their ideas on this topic.

**PN 08 How Can Evaluation Help Bureaucracies Be, Well, Less Bureaucratic?****PC 031 - The Disconnect Between Outcome-Oriented Public Services and Activity-Based Bureaucratic Practices**

*B. Perrin<sup>1</sup>*

<sup>1</sup> *Independent consultant, independent, Vissec, France*

Burt will distinguish between bureaucracy as an organisational form or structure essential to democracy, vs. bureaucratic practices that may or may not always serve their intended purpose. The importance of an outcome focus to public services is now generally recognised. But bureaucratic practices, and assessments, tend to be based upon compliance with rules and procedures, with activities rather than with outcomes. How can this apparent conflict be resolved? Bureaucratic practices are often viewed as “undiscussable” – but surely should be open to empirical assessment through evaluation.

**PN 08 How Can Evaluation Help Bureaucracies Be, Well, Less Bureaucratic?****PC 032 - Management Paradigms, Accountability, Evaluation, and Bureaucracy in (Un)Predictable Contexts**

*J. Toulemonde<sup>1</sup>*

<sup>1</sup> *Independent evaluation expert, independent, Journans, France*

Jacques will discuss implications of uncertainty and complexity for evaluation. He will indicate how many typical accountability and bureaucratic processes may fail to take into account this context, and thus be counterproductive in nature.

**PN 08 How Can Evaluation Help Bureaucracies Be, Well, Less Bureaucratic?****PC 033 - Evaluation Quality Assessment: Useful Tools vs. Bureaucratic Impediments to Learning?**

*A. Pollard<sup>1</sup>*

<sup>1</sup> *DFID, Evaluation Unit, London, United Kingdom*

Alison will consider various evaluation quality assessment approaches. Under what circumstances can these represent useful tools to aid in improving the quality of evals? Or, when might they serve instead as bureaucratic impediments to learning?

**PN 08 How Can Evaluation Help Bureaucracies Be, Well, Less Bureaucratic?**

## **PC 034 - Panel Contribution**

F. Leeuw<sup>1</sup>

<sup>1</sup> University of Maastricht, Faculty of Law, Dept of Methods and Foundations, Maastricht, The Netherlands

**PN 08 How Can Evaluation Help Bureaucracies Be, Well, Less Bureaucratic?**

## **PC 035 - Evaluative Thinking and Bureaucracy**

P. Hawkins<sup>1</sup>

<sup>1</sup> Creative Evaluation Limited, CEO, Wellington, New Zealand

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 09 Building Strong Monitoring and Evaluation Systems for Global Partnership Programs**

## Chairing the Session on “Building Strong Monitoring and Evaluation Systems for Global Partnership Programs”

A. Aghumian<sup>1</sup>, E. Di Gropello<sup>1</sup>, N. Khattri<sup>2</sup>, J. Garcia<sup>3</sup>, K. Johnson<sup>4</sup>

<sup>1</sup> World Bank Group, Independent Evaluation Group, Washington- DC, USA

<sup>2</sup> Global Partnership for Education, Global Partnership for Education, Washington- DC, USA

<sup>3</sup> Global Environment Facility, Independent Evaluation Office, Washington- DC, USA

<sup>4</sup> World Bank Group, Evaluation Capacity Development- Independent Evaluation Group, Washington- DC, USA

Session Title: Building strong monitoring and evaluation systems for global partnership programs. Global partnership programs that involve many stakeholders have become important vehicles for coordinating work on global priorities, providing global, regional, and national public goods, responding to crises and fragility. Their role is growing in the age of the SDGs, with many sector-specific targets and need to tap in more public and private financing. Building credible evidence on the development effectiveness of partnership programs and on how well they integrate into the global aid architecture is therefore a priority. These programs often strive to produce good results reporting and demonstrate value for money. Many programs invest a good deal in monitoring and evaluation, and commission regular external evaluations. Yet, building an effective and efficient M&E system is still a challenge for many. How do these multi-stakeholder and multi-layer programs build and use their M&E systems? How they can ensure that the evidence from the M&E is useful and feed into decision-making? How can the entire M&E process be kept cost-efficient and streamlined despite multiple stakeholder interests, and despite operating in fragile or post-crisis situations often require tailored approaches? What value can an independent evaluation add on top of this? While the research on monitoring and evaluation in development has become quite advanced to help understand the complexity of contexts in which these programs operate, the practical application is still challenging due to the evolving nature of partnership programs. The panel session will bring together the M&E specialists of three global partnership programs in environment, education and evaluation capacity building (the Global Environment Facility, the Global Partnership for Education, and the CLEAR Initiative (the Regional Centers for Learning on Evaluation and Results) and the Independent Evaluation Group's cross-cutting perspective stemming from systematic evaluation of global partnership programs in different sectors in the last decade. The objective of the panel is to share key monitoring and evaluation challenges the global partnership programs face, to discuss different approaches and perspectives of the programs, and to share some innovative approaches and solutions to those M&E problems in different settings. The panel expects to use this session to engage with evaluators and M&E specialists, evaluation commissioners and users from other development agencies in discussion of the role of internal M&E systems, role of independent evaluations.

**PN 09 Building Strong Monitoring and Evaluation Systems for Global Partnership Programs**

## PC 036 - Independent Evaluation Group

A. Aghumian<sup>1</sup>

<sup>1</sup> World Bank Group, Independent Evaluation Group, Washington, USA

The presentation will share IEG's experience in evaluating global multi-stakeholder partnerships with the focus on cross-cutting lessons on building effective M&E and better use of external evaluations.

**PN 09 Building Strong Monitoring and Evaluation Systems for Global Partnership Programs**

**PC 037 - Global Partnership for Education (GPE) Experience**

*N. Khattri*<sup>1</sup>

<sup>1</sup> *Global Partnership for Education, Global Partnership for Education, Washington- DC, USA*

The presentation will focus on the approaches for, and challenges in, evaluating a partnership program that focuses on country-wide interventions.

**PN 09 Building Strong Monitoring and Evaluation Systems for Global Partnership Programs**

**PC 038 - Global Environment Facility (GEF) Experience**

*J. Uitto*<sup>1</sup>

<sup>1</sup> *GEF Independent Evaluation Office, Director, Washington DC, USA*

The presentation will be on a framework for long-term, cost-effective monitoring that links global, national and local data sources at the Global Environment Facility.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 10 The Evaluation Marketplace: Exploring the Market Dynamics of Evaluation**

## The Evaluation Marketplace: Exploring the Market Dynamics of Evaluation

*C. Christie*<sup>1</sup>

<sup>1</sup> *UCLA, Education, Los Angeles, USA*

The motivation for this panel emerges from an apparent paradox: Despite the fact that most evaluation practice is deeply embedded in a broader market of evaluative knowledge production, complete with contractual negotiations and obligations, competing providers and consumers, this fundamental aspect of evaluation has received scant scholarly attention. If evaluation is intrinsically embedded within these market dynamics of commissioned – and in effect commercial – knowledge production, then examining the extent to which and how these dynamics influence evaluation practice becomes topical. The proposed panel aims to describe and discuss the marketplace of evaluation in a select set of countries. More specifically, we intend to tease apart the various ways in which the marketplace shapes the practice and profession of evaluation. The vast majority of evaluative practice is formed, shaped, and grounded upon market dynamics. By not awarding attention to the ways in which the scope, design, methodology and deliverables are shaped by the push and pull of market players, we fail to appreciate the fundamental conditions for the evolution of evaluation as a field of practice. The panel will contribute to the field of evaluation in at least three ways (1) by presenting a framework for better understanding evaluation markets, (2) by providing a systematic empirical examination of the market forces and dynamics that influence evaluation practice, and (3) by paving and furthering the way for future research on this important topic.

**PN 10 The Evaluation Marketplace: Exploring the Market Dynamics of Evaluation**

## PC 039 - The Evaluation Marketplace – A Conceptual Framework

*S. Bohni-Nielsen*<sup>1</sup>

<sup>1</sup> *Municipality of Gribskov- Denmark, Education, Gribskov, Denmark*

This presentation provides the conceptual foundation on which the remainder of the panel is grounded. First, the presentation considers the idea of an evaluation marketplace and reflects on the limited attention awarded the commercial aspects of evaluation. Second, the presentation presents the Evaluation Market Framework, an emerging framework for better understanding the composition, context, and dynamics of evaluation markets.

**PN 10 The Evaluation Marketplace: Exploring the Market Dynamics of Evaluation**

## PC 040 - The Evolving Market for Systematic Evaluation in Canada

*R. Lahey*<sup>1</sup>, *C. Elliott*<sup>2</sup>, *S. Heath*<sup>3</sup>

<sup>1</sup> *REL Solutions Inc., Evaluation, Ottawa, Canada*

<sup>2</sup> *University of Ottawa, School of Management, Ottawa, Canada*

<sup>3</sup> *University of Ottawa, Education, Ottawa, Canada*

The presentation describes the market for evaluation in Canada, both the supply side and the demand side, identifying the characteristics of the structure of the industry, its evolution over the last 40 years, as well as market behavior of key components of the industry.

**PN 10 The Evaluation Marketplace: Exploring the Market Dynamics of Evaluation****PC 041 - Does Reputational Capital Matter within Italy's Quasi-Market for Evaluation?**

M. Mara<sup>1</sup>

<sup>1</sup> University of Salerno, Public Policy, Salerno, Italy

In this paper, I examine Italy's quasi market for evaluation highlighting such issues as the structure of the marketplace; the economic, financial, organizational, regulatory, and political barriers to entry; the value of evaluative services in relation to management consulting services; and last but not least the educational and professional pathways for practitioners' accreditation. I then discuss some ethical issues related to cooperation between commissioners and experts and the conflict of personal and professional interests among experts.

**PN 10 The Evaluation Marketplace: Exploring the Market Dynamics of Evaluation****PC 042 - The US Federal Evaluation Market**

C. Christie<sup>1</sup>

<sup>1</sup> UCLA, Education, Los Angeles, USA

This presentation describes the main features of the demand side of the U.S. federal evaluation market. First, the presentation provides a brief historical account of the overarching developments of the U.S. federal evaluation market, awarding special attention to current trends in federal evaluation funding. The second part of the presentation offers a more fine-grained analysis of evaluation procurement and contracting processes in the Department of Health and Human Services.

**PN 10 The Evaluation Marketplace: Exploring the Market Dynamics of Evaluation****PC 043 - The Commercial of Evaluation: A Danish Case Study**

S. Lemire<sup>1</sup>

<sup>1</sup> UCLA, Education, Los Angeles, USA

The final presentation examines the Danish evaluation market and how market forces affect the evolution of an evaluation enterprise. These forces include: buyer power, supplier power, competitive rivalry, threat of substitution, and threat of new entry.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

## PN 11 Evaluation for Improving Public Sector Efficiency

### Evaluation for Improving Public Sector Efficiency

*G. Korella*<sup>1</sup>

<sup>1</sup> *Individual Expert, Individual Expert in Public Administration, Athens, Greece*

**Rationale:** Performance is a key indicator in the reform process of the public sector. However, although governments engage in institutional reforms to improve public sector performance, there is little documentation on the impact of these reforms on efficiency.

Public sector decision makers face a number of challenges when introducing performance information as a tool for improving efficiency: design clear targets at the planning phase; regulate the necessary framework; identify relevant methodology and tools; select valid impact indicators; create benchmarking; ensure adequate resources to conduct evaluations; acquire trust by all stakeholders;

Although evaluation is strongly encouraged by the EU, as an effective tool for the enhancement of performance of the public sector, the vast majority of public institutions do not use evaluation as a tool for improvement. Often, there is no structured approach of evaluation in the public administration, poor benchmarking and comparative analyses, lack of understanding and capacity.

**Objectives:** The objective of the proposed panel is to advocate on the positive results of evaluation in enhancing efficiency, effectiveness and economy of public services.

Discussion will focus on:

- the benefits of evaluation in running public services more efficient
- the challenges of using evaluation as a tool to run public services more efficient
- the methodology of using evaluation as a tool to run public services more efficient
- the ways to communicate the close link between evaluation and efficiency to all stakeholders, including final users of public services

The areas on which the members of the panel shall present their views and share their experience are, inter alia:

- Which basic measurements, indicators and benchmarking would lead – if applied – to more efficient public services
- How to design the evaluation framework in order to influence efficiency of a public entity
- How to establish an evaluation – friendly framework which will lead to a more efficient public sector
- How to overcome common barriers and resistance to change
- Ways to benefit from best practices, networks and open data.

**Brief justification:** The proposed panel will increase awareness on the tight links between assessing performance and enhancing efficiency. It is expected to contribute to the ongoing discussion on how to improve capacity of the public sector, through innovative processes and applied tools, related to performance measuring.

**PN 11 Evaluation for Improving Public Sector Efficiency****PC 044 - Public Management by Objectives**

G. Theodorakis<sup>1</sup>

<sup>1</sup> Ministry of Administrative Reconstruction, Secretary General, Athens, Greece

**PN 11 Evaluation for Improving Public Sector Efficiency****PC 045 - How to Measure Good Governance: The Twelve European Principles of Good Democratic Governance**

A. Tatarenko<sup>1</sup>

<sup>1</sup> Council of Europe, Head of the Centre of Expertise for Local Government Reform, Strasbourg, France

The **12 Principles** of Good Democratic Governance were endorsed by a HYPERLINK "[https://search.coe.int/cm/Pages/result\\_details.aspx?ObjectID=09000016805d3dc8](https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016805d3dc8)" decision of the Committee of Ministers of the Council of Europe in 2008. They represent a common vision of Good Governance in Europe.

The Centre of Expertise has developed **a benchmark** which allows to measure the quality of governance against these 12 principles, at any level of government.

The **European Label of Governance' Excellence (EloGE)** is awarded to local authorities having achieved a high overall level of governance measured against this benchmark.

Local authorities who wish to apply for the Label are evaluated according to the following tools: a matrix evaluating their actions, a questionnaire addressed to citizens, and a questionnaire addressed to local elected representatives.

This evaluation and improvement tool allows local authorities to understand their strengths and their weaknesses when providing public services and exercising public authority. To improve the quality of local governance, local authorities can also use the various tools developed by the Centre of expertise and take inspiration from their colleagues' best practices.

More information is available here: <https://www.coe.int/en/web/good-governance/12-principles-and-elope>

**PN 11 Evaluation for Improving Public Sector Efficiency****PC 046 - Panel Contribution**

G. Korella<sup>1,2</sup>

<sup>1</sup> Individual Expert, Individual Expert in Public Administration, Athens, Greece

<sup>2</sup> European Evaluation Society (member), Hellenic Evaluation Society (member), administrative reforms, Athens, Greece

Wednesday, 3 October 2018  
12:00–13:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 12 Powers and Privileges: The Non-Visible Side of Evaluations. Lessons From Latin America**

## **Powers and Privileges: The Non-Visible Side of Evaluations. Lessons From Latin America**

R. Luna<sup>1</sup>, M. Tarsilla<sup>2</sup>, F. Amariles<sup>3</sup>, S. Salinas<sup>4</sup>

<sup>1</sup> University of Cartagena, Ipreg, Cartagena de Indias, Colombia

<sup>2</sup> UNICEF, West and Central Africa, Dakar, Senegal

<sup>3</sup> Learning for Impact / Universidad del Valle, Evaluation, Weston- FL, USA

<sup>4</sup> Independent consultant / REDMEBOL, President, La Paz, Plurinational State of Bolivia

Evaluations involve interactions between people, and imply power relations, an aspect seldom addressed in evaluations. Also, evaluations influence – by action or omission – pre-existing hierarchies and power relations in the context of the evaluation process; in this regard, they can either exacerbate or challenge and contribute to transform those relationships that reproduce gender inequalities and other injustices of the existing social order. Evaluators can contribute to revert unequal power relations existing in the contexts and among the subjects of the evaluation. This requires assuming one's own position and subjectivity, recognizing ourselves as transforming actors and developing tools and attitudes for transformative evaluation. We will share our collaborative learning in LAC contexts about approaches and tools to address the critical knots of power in evaluations: Breaking the power status quo; transforming hegemonic masculinity from a feminist perspective; young evaluators challenging adultism in evaluation; breaking paradigms and the colonial heritage of knowledge in the rural sector. The discussion about power in evaluations is a key dimension, which on the one hand alludes and can challenge ethical but even epistemological aspects of the findings, including the rigor and validity of data in situations when the voices of the most socially excluded are not included in the evaluation process. This is even more so under the SDG principle of “no one left behind”. On the other hand, it puts on the agenda the role of evaluation regarding the status quo and, specifically, the positioning regarding the situation of inequality, exclusion and marginalization of women and other groups in most of the contexts where evaluations are carried out, and the potential of each evaluator as an actor of change helping to build the future from the evaluation process. From our experiences as evaluators, we have learned that we arrive with a social and historical background to the contexts studied and that this inevitably influences our positioning and perspectives, and how others perceive and interpret our presence, the explicit and implicit expectations of the process and how different actors interact with us. Consequently, in the interactions with the groups and sub-groups evaluated, we often reproduce colonial, patriarchal and adultist relationships in a generally uncritical and unreflective manner.

Our bias is translated and reflected in different aspects and scopes of the evaluation, which have to do with the conception of the “otherness” and the forms of relationship that derive from it, with the approach to the context, with the methodologies and procedures put in place, with the inclusions and exclusions of the process, with the indicators, as well as in the definition and application of the ethical frameworks and norms. This affects the process and the results of the evaluation, even questioning central values of “good evaluations” like rigor and objectivity. Presentations from this panel will help collectively build new approaches, methodologies, and evaluative tools to address the structural problems that perpetuate inequalities. Through making visible and addressing inequalities, the field of evaluation creates bridges to advocacy for policies and programs that address the problems identified.

**PN 12 Powers and Privileges: The Non-Visible Side of Evaluations. Lessons From Latin America****PC 050 - Young and Emerging Evaluators (YEEs): Experiences, Challenges and Opportunities in the Face of Inclusion with Equity**

*R. Luna<sup>1</sup>, M. Tarsilla<sup>2</sup>, F. Amariles<sup>3</sup>, S. Salinas<sup>4</sup>*

<sup>1</sup> *University of Cartagena, Ipreg, Cartagena de Indias, Colombia*

<sup>2</sup> *UNICEF, West and Central Africa, Dakar, Senegal*

<sup>3</sup> *Learning for Impact / Universidad del Valle, Evaluation, Weston- FL, USA*

<sup>4</sup> *Independent consultant / REDMEBOL, President, La Paz, Plurinational State of Bolivia*

The objective of this presentation is to discuss the challenges and opportunities faced by YEEs in the early stages of the evaluation practice through the experiences and perspectives of the author, as well as based on previous studies, so that finally we can all contribute to answer the question: How can we advocate and include the voices of young people in the evaluation process? To this end, challenges such as: the traditionalist approach in evaluation, difficulty in entering the labor market, the antagonism between the young and adult perspective, lack of information on opportunities, the incipient offer of training in the LAC region, the adultism in evaluation teams, the lack of adequate methodologies to evaluate the young population, among others. Important opportunities for the YEEs are emerging, such as innovation, dynamism, and extensive use of ICTs, among others, which ensure their effective participation in the discipline of evaluation.

**PN 12 Powers and Privileges: The Non-Visible Side of Evaluations. Lessons From Latin America****PC 051 - Power and Gender Issues: Reflections and Proposals From the Field of Evaluation**

*R. Luna<sup>1</sup>, M. Tarsilla<sup>2</sup>, F. Amariles<sup>3</sup>, S. Salinas<sup>4</sup>*

<sup>1</sup> *University of Cartagena, IPREG, Cartagena de Indias, Colombia*

<sup>2</sup> *UNICEF, West and Central Africa, Dakar, Senegal*

<sup>3</sup> *Learning for Impact / Universidad del Valle, Evaluation, Weston- FL, USA*

<sup>4</sup> *Independent consultant / ReLAC and REDMEBOL's President, Evaluation, La Paz, Plurinational State of Bolivia*

The approach to existing hierarchies and inequalities in evaluation contexts is rarely addressed as an important issue of power relations that goes beyond methodological matters – e.g. How to organize an interview with women. At the same time, as a response to the new paradigms that call for decolonization in order to address inequalities, there are still voices that claim “cultural respect”, questioning any attempt to “interfere” in internal affairs, and at the same time legitimizing the established order and culturally embedded inequalities. From the point of view of evaluations, the difference and the decision is on whether to carry out an evaluation “with gender sensitivity” or if it is to contribute to transform unequal gender relations? At the end of the evaluation is it to verify if the project “did what it said it was going to do”, or if what it did contributed to social change? And what is my position as an evaluator in the face of gender inequalities that result from power hierarchies?

**PN 12 Powers and Privileges: The Non-Visible Side of Evaluations. Lessons From Latin America****PC 052 - New Approaches to Evaluate from the Feminist Perspective: Towards a Non-Hegemonic Masculinity in Evaluation**

R. Luna<sup>1</sup>, M. Tarsilla<sup>2</sup>, F. Amariles<sup>3</sup>, S. Salinas<sup>4</sup>

<sup>1</sup> University of Cartagena, IPREG, Cartagena de Indias, Colombia

<sup>2</sup> UNICEF, West and Central Africa, Dakar, Senegal

<sup>3</sup> Learning for Impact / Universidad del Valle, Evaluation, Weston- FL, USA

<sup>4</sup> Independent consultant / REDMEBOL, President, La Paz, Plurinational State of Bolivia

The synergy between the new masculinities approach and the principles of feminist evaluation is shown. This presentation describes the close links that the new masculinities approach has developed with feminist theory and politics, its strategy to work on the personal change of men towards a non-hegemonic masculinity. Examples of the proposed actions with men for equality that have been included in projects and cooperation programs in various institutions in Latin America are provided. Some methodologies are outlined to work on an approach of new masculinities for projects, programs and public policies of international cooperation. Good practices in evaluation are described by combining the principles of feminist evaluation with the approach of new masculinities; Evaluation methods are explored to measure the changes in men towards more egalitarian positions that have an important potential for social change and human development, and a summary of results from a real evaluation that applied these methods is presented.

Wednesday, 3 October 2018  
12:00–13:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 13 Measuring the Impact of Structural Reforms carried out in Times of Crisis**

## Measuring the Impact of Structural Reforms carried out in Times of Crisis

*D. Ioannou<sup>1</sup>*

<sup>1</sup> Hellenic Evaluation Society, President, Athens, Greece

<sup>2</sup> European Evaluation Society, Member

**Rationale:** In the recent years many countries have undergone intensive structural reforms with the aim to restructure their socio-economic development base and tackle some of the underlying causes and conditions of the crisis. There is a shared recognition among policymakers that changes in the economic-governmental structure alone cannot restore a system of fundamental stimulus for growth. A range of important direct or indirect effects which go beyond the regulatory, institutional or political change may be more important than the short-term results of the reform measures. Shedding light on the mechanisms of the structural change and on the long-term impact of the reforms require more in-depth analytical approaches, especially in the frame of unpredictable context.

**Objectives sought:** The objective of the proposed panel is to discuss: – The role of evaluations in assessing the long-term effects of structural reforms for the society – Evaluation approaches to capture the effects of structural reforms under different growth models – Structural reforms and societal change – capturing the pressures on the environment in transition.

**Justification:** This topic is of particular relevance to the topic of the conference and to members of the evaluation community, public sector officials involved in evaluation, to the academia and to the wider public. It is expected to generate a debate on the role of evaluation in evidencing areas of adjustment to stimulate the long term sustainable effects of structural reforms and shaping future restructuring practice. It will also trigger discussion on the qualifications and skills of the evaluators relative to the assessment of the impact of the reforms in the post global-crisis period.

**PN 13 Measuring the Impact of Structural Reforms carried out in Times of Crisis**

**PC 053 - Panel Contribution**

*D. Ioannou*<sup>1</sup>

<sup>1</sup> Hellenic Evaluation Society, President, Athens, Greece

<sup>2</sup> European Evaluation Society, Member

**PN 13 Measuring the Impact of Structural Reforms carried out in Times of Crisis**

**PC 054 - Panel Contribution**

*D. Dotto*<sup>1</sup>

<sup>1</sup> European Commission, Secretariat General, Structural Reforms Support Service, Brussels, Belgium

**PN 13 Measuring the Impact of Structural Reforms carried out in Times of Crisis**

**PC 055 - Panel Contribution**

*G. Mergos*<sup>1</sup>

<sup>1</sup> Professor Emeritus- University of Athens, Department of Economics, Athens, Greece

**PN 13 Measuring the Impact of Structural Reforms carried out in Times of Crisis**

**PC 056 - Panel Contribution**

*B. de Laat*<sup>1</sup>

<sup>1</sup> European Evaluation Society, Board, Paris, France

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

## PN 14 Looking to the Criticality of Oversight and the Necessity of Working Together

### Looking to the Criticality of Oversight and the Necessity of Working Together

*T. Schwandt*<sup>1</sup>

<sup>1</sup> University of Illinois, Indianapolis, USA

Historically, evaluation units have worked isolated, implementing their own strategy and working plans, with interaction limited mostly to a few joint evaluations. Nevertheless, over the last decade, international organizations have increasingly been under pressure to join forces and work together to develop synergies, rationalize resources and produce more coherent results.

In a context of decreasing resources, growing investment needs, and higher transparency and accountability demands, organizations – and their oversight units – must explore working jointly and across traditional professional boundaries. This is particularly important now that the complexity and broadness of the sustainable development goals require more than ever cohesive and coordinated interventions to support countries in their efforts to build more resilient societies and achieve sustainable development.

Through audits and performance evaluations, the oversight function ensures that development resources are wisely spent and expected results are achieved in a sustainable manner. It also provides evaluative evidence of the organization's performance to learn and inform policy decisions.

There are valuable lessons that can be drawn from some joint initiatives, such as the evaluation of UNDP institutional effectiveness conducted by the Independent Evaluation Office (IEO) and the Office of Audit and Investigations (OAI) of this organization.

Though value for money is becoming more critical with the decrease in funding, the oversight function has to look beyond the efficiency component to look also at effectiveness, otherwise it can undermine the development agenda.

The objectives of the panel will be to discuss challenges and opportunities for joint evaluations and reach a better understanding of how additional criteria or innovative approaches can help to assess complex multi-stakeholder interventions.

The panel discussion should be answering the following questions:

- What are the challenges and opportunities of following-up and evaluating multi-stakeholder initiatives?
- What lessons have we learnt when overseeing jointly organizations' performance?
- How can those lessons be applied in unstable situations?
- How can oversight and evaluation units maintain their independence, relevance and responsiveness capacity while responding to the imperative to work in a more coherent and cross-cutting way?
- How can the DAC criteria be revisited to accommodate joint oversight activity?
- How to ensure that the oversight function contributes to build resilience to climate changes, natural disasters, security and many other threats that endanger human well-being, and especially vulnerable populations?

The panel presentations and subsequent debate will:

- Deepen the debate on the importance of meeting common high-quality norms and standards.
- Answer some of the questions that have arisen lately in the evaluation community due to the complexity of the 2030 Agenda, including how to contribute to advance the public interest.
- Explore innovative approaches to evaluation to build more resilient societies.

**PN 14 Looking to the Criticality of Oversight and the Necessity of Working Together****PC 058 - Audit and Evaluation: Working Collaboratively to Support Accountability**

*I. Naidoo*<sup>1</sup>

<sup>1</sup> *United Nations Development Programme UNDP, Independent Evaluation Office, New York, USA*

There have been growing calls for oversight units to work in a more collaborative manner, with arguments presented by the advocates that this avoids duplication and promotes synergies.

In the UN system, and UNDP in particular, this has ignited a conversation between evaluation and audit offices. Whilst there are many benefits to be gained from collaboration, without compromising professional identity and coherence, the practice requires strong leadership, in particular from the heads of collaborating units.

In a 1<sup>st</sup> ever joint evaluation of the Institutional Effectiveness of UNDP (2017), the resultant report was well received by the Executive Board of the organization, who found that the assessment helped clarify concepts, in particular about efficiency (the traditional purview of auditors) and effectiveness (evaluators), and that it was possible to understand performance more as a gestalt, that two quite different organizational streams.

The report forged a common language, through the negotiations between evaluators and auditors' language, that was critical, and an even more compelling reason for collaboration was the discussion on "evidence". This term has become quite amorphous, and defining more explicitly what constitutes evidence in the evidence hierarchy, how quantitative and qualitative data is processed and valued, and what triangulation and mixed methods mean, advanced both offices; their capacity, knowledge and visibility.

This paper talks to the strengths and challenges of the collaboration and offers some learnings, for what is likely to become a more common undertaking in the UN oversight architecture.

**PN 14 Looking to the Criticality of Oversight and the Necessity of Working Together****PC 060 - Evaluation 'Architectures' – Implications of How we Work Together**

*E. Stern*<sup>1</sup>

<sup>1</sup> *Lancaster University- UK, Department: Educational Research, Lancaster, United Kingdom*

Evaluations of complex initiatives need to reflect the diversity of what they evaluate whilst also striving for an appropriate degree of coherence in their conclusions. Evaluations are located within an 'architecture' – institutional arrangements of planning, funding, support, implementation and reporting. Coordination mechanisms such as Independent Evaluation Offices; Joint evaluations; and cross-agency evaluation cooperation and funding; contracting and feedback arrangements; opportunities for peer-learning are all examples of evaluation architectures. These architectures instantiate assumptions about how best to evaluate policies and programmes, have methodological implications and can in some circumstances even delimit the kinds of results an evaluation might produce. This input will consider the implications of different evaluation architectures for notions of institutional effectiveness and efficiency on the one hand; and wellbeing, resilience and sustainability on the other.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

#### PN 15 Privacy, GDPR and the Responsible Use of Data in Evaluation Systems

### Privacy, GDPR and the Responsible Use of Data in Evaluation Systems

M. McGuire<sup>1</sup>, G. Kerr<sup>2</sup>

<sup>1</sup> LogicalOutcomes, Evaluation, Mississauga, Canada

<sup>2</sup> LogicalOutcomes, President Emeritus and Founder, Toronto, Canada

**Rationale:** With the new European Union General Data Protection Regulation (GDPR) rules coming into force in May 2018, many evaluators may be unsure of how to comply with these new rules. This session will discuss the regulations, specific processes for complying and integration with other rules required by international organizations.

**Objectives:** By the end of this session, participants should have an understanding of how GDPR may affect their evaluations, its impact on data collection/storage, examples of procedures for complying with the rules and introduction to a system with security procedures that other organizations can use in their own systems.

**Narrative:** Any evaluations that collect personal information from EU residents or citizens, or that are carried out by organizations located in the EU must comply with strict privacy requirements. Personal information includes things like IP addresses and political opinions, not only names and identity numbers. This creates another level of complexity when conducting evaluations. At the same time it supports evaluations that are conducted with respect for human dignity and individual rights. The new rules create additional challenges in the design and management of evaluations. How are evaluators dealing with GDPR, and how can they prepare for it while keeping costs down? The Service Information System (SIS), a Canadian-based monitoring and evaluation platform launched in 2018, is designed to be compliant with GDPR. By building SIS on open source software and client-managed encryption keys on top of Microsoft Office 365, we developed an inexpensive platform with security procedures that other organizations can use in their own systems. We hope to work with the international community to develop and share open tools that will protect client privacy in the context of low IT capability, constrained resources and often hostile environments. We will share the challenges and how they have been addressed through a platform with security procedures.

#### PN 15 Privacy, GDPR and the Responsible Use of Data in Evaluation Systems

### PC 061 - Designing Monitoring and Evaluation Systems for Resilient Societies

M. McGuire<sup>1</sup>, G. Kerr<sup>2</sup>

<sup>1</sup> LogicalOutcomes, Evaluation, Mississauga, Canada

<sup>2</sup> LogicalOutcomes, President Emeritus and Founder of LogicalOutcomes, Toronto, Canada

**Rationale:** Increasingly all organizations are expected to monitor progress on a regular basis and conduct evaluations. Many organizations find it challenging to meet these expectations, partly due to the expense and complexity of setting up monitoring and evaluation systems. They need less costly credible approaches and methods.

**Objectives:** By the end of this session participants will have a better understanding of the role of standard definitions of indicators and measures that will support their evaluations and reduce cost while at the same time increasing the impact of their findings. This presentation will provide ideas on how evaluators can support all organizations in improved monitoring and evaluation. They will be introduced to the Service Information System (SIS), an open source evaluation platform for NGOs and its use with open source online data collection and storage systems such as DHIS2.

**Narrative:** The panel will present ideas for rethinking evaluation methods and design, looking at indicator development, data collection, analysis and reporting issues in order to decrease unpredictability to assist organizations in improving people's lives and making the organizations more resilient. We will explore: Reducing the costs of monitoring and evaluation systems by using open tools, including setup and design, validated indicators, security, training, the responsible use of data, the burden of data collection for agencies and respondents, and the creation of meaningful and useful reports; We will introduce SIS which includes an indicator registry and community data portal. The importance of shared indicators using international metadata standards that are tied to the Sustainable Development Goals. This dramatically reduces the cost and improves the validity of evaluation while enabling evaluators to aggregate information across sectors and organizations to look for promising approaches; and. The benefits of open source software in terms of community control over health information systems rather than giving ownership of data and systems to for-profit companies. Also the opportunity of building open source tools as a community, including indicators, methodologies and instruments, rather than proprietary locked-in systems.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 16 Outcome Harvesting Customized for Monitoring and Evaluation in the Dutch Dialogue & Dissent Lobbying and Advocacy Programme for Development**

## Outcome Harvesting Customized for Monitoring and Evaluation in the Dutch Dialogue & Dissent Lobbying and Advocacy Programme for Development

*W. Richert<sup>1</sup>, K. Chambille<sup>2</sup>, K. Biesbrouck<sup>3</sup>, N. Van der Jagt<sup>4</sup>*

<sup>1</sup> *Wolfgang Richert Consulting, Independent Consultant, Amsterdam, Netherlands*

<sup>2</sup> *Hivos, Programme quality, The Hague, Netherlands*

<sup>3</sup> *Oxfam Novib, Impact measurement and knowledge team, The Hague, Netherlands*

<sup>4</sup> *The Netherlands Institute for Multiparty Democracy NIMD, Planning- Monitoring and Evaluation Coordinator, The Hague, Netherlands*

The Dutch Ministry of Foreign Affairs supports a major Dialogue and Dissent programme to support lobbying and advocacy to influence development around the world. In this panel we present challenges in using Outcome Harvesting in this programme and then three cases of Dutch organisations highlight how they customised Outcome Harvesting.

Ricardo Wilson-Grau, whose book *Outcome Harvesting for Monitoring and Evaluation: – Practical Applications of Essential Principles* will be published by IAP to coincide with the EES 2018 Conference, will chair the session. Wolfgang Richert, an independent evaluator based in the Netherlands, was one of the evaluators in the seminal evaluation that brought Outcome Harvesting to the attention of the Dutch Ministry of Foreign Relations. Wolfgang supports the use of Outcome Harvesting in several consortia involved in the Dialogue and Dissent (D&D) program. He will explain the M&E challenges faced by the 25 consortia involved in the D&D program and the potential of Outcome Harvesting to meet those challenges. This introduction is followed by examples of customising the introduction and use of the methodology. Karel Chambille is a senior advisor on programme quality at Hivos, a Dutch development cooperation agency. With special responsibility for monitoring, evaluation and learning, Karel will share the experience of introducing Outcome Harvesting at Hivos as a monitoring approach for the D&D programmes. The presentation will reflect on Hivos' experience of introducing Outcome Harvesting as an approach for monitoring results of a global programme of Lobby and Advocacy, and will cover aspects of capacity development of programme staff (in Hivos and partner organisations) as well as aspects of use of the 'harvested outcomes' for learning and accountability. Karen Biesbrouck is the project leader on Outcome Harvesting for the impact measurement and knowledge team of Oxfam Novib. With Sanne Djojoseparto, a monitoring, evaluation and learning colleague at Oxfam Novib, Karen will explain how Oxfam Novib applies outcome harvesting in monitoring a complex D&D program (31 projects in 17 countries, each project being an alliance with several local partners). Series of outcome harvests resulted in a wealth of descriptions of observable social change and contributions to these changes, including Oxfam's and its partners' contributions. In addition to each project being able to draw contextual lessons learnt, Oxfam Novib's large-scale deductive analyses of qualitative data in all outcome statements enables assessment and validation of the theory of change. The results of the analysis in the Right to Food sub-programme will be presented. Nic Van der Jagt is the planning, monitoring and evaluation coordinator for The Netherlands Institute for Multiparty Democracy (NIMD). Nic will explain why performance evaluation using DAC criteria does not always work and the lessons learned from a pilot with the Outcome Harvesting alternative in the form of a mid-term evaluation of four NIMD D&D country programmes. NIMD is a democracy assistance organization that supports political parties in developing democracies through interparty dialogue whose effectiveness is especially difficult to measure with classical performance evaluations. He will present the experiences with an Outcome Harvesting in Mali and Mozambique, describing the extent to which the outcomes achieved represent patterns of progress towards respective programme objectives. Attention will be paid to the substantiation step of the OH methodology: How to determine if the outcome information is credible enough for learning and accountability.

**PN 16 Outcome Harvesting Customized for Monitoring and Evaluation in the Dutch Dialogue & Dissent Lobbying and Advocacy Programme for Development****PC 062 - The Challenges of the use of Outcome Harvesting for Consortia in the Dutch Dialogue and Dissent (D&D) Development Program**

W. Richert<sup>1</sup>

<sup>1</sup> Wolfgang Richert Consulting, Independent evaluator, Amsterdam, Netherlands

Wolfgang Richert, an independent evaluator based in the Netherlands, was one of the evaluators in the seminal evaluation that brought Outcome Harvesting to the attention of the Dutch Ministry of Foreign Relations. Wolfgang supports the use of Outcome Harvesting in several consortia involved in the Dialogue and Dissent (D&D) program. He will explain the M&E challenges faced by the 25 consortia involved in the D&D program and the potential of Outcome Harvesting to meet those challenges.

**PN 16 Outcome Harvesting Customized for Monitoring and Evaluation in the Dutch Dialogue & Dissent Lobbying and Advocacy Programme for Development****PC 063 - Using Outcome Harvesting as an Approach for Monitoring Lobby & Advocacy Programs**

K. Chambille<sup>1</sup>

<sup>1</sup> Hivos, Programme quality, The Hague, Netherlands

Karel Chambille is a senior advisor on programme quality at Hivos, a Dutch development cooperation agency. With special responsibility for monitoring, evaluation and learning, Karel will share the experience of introducing Outcome Harvesting at Hivos as a monitoring approach for the D&D programmes. The presentation will reflect on Hivos' experience of introducing Outcome Harvesting as an approach for monitoring results of a global programme of Lobby and Advocacy, and will cover aspects of capacity development of programme staff (in Hivos and partner organisations) as well as aspects of use of the 'harvested outcomes' for learning and accountability.

**PN 16 Outcome Harvesting Customized for Monitoring and Evaluation in the Dutch Dialogue & Dissent Lobbying and Advocacy Programme for Development****PC 064 - Outcome Harvesting: Oxfam Novib's Large-Scale Deductive Analysis of Outcome Statements Validating a Theory of Change**

K. Biesbrouck<sup>1</sup>

<sup>1</sup> Oxfam Novib, Impact measurement and knowledge team, The Hague, Netherlands

Karen Biesbrouck is the project leader on Outcome Harvesting for the impact measurement and knowledge team of Oxfam Novib. With Sanne Djojoseparto, a monitoring, evaluation and learning colleague at Oxfam Novib, Karen will explain how Oxfam Novib applies outcome harvesting in monitoring a complex D&D program (31 projects in 17 countries, each project being an alliance with several local partners). Series of outcome harvests resulted in a wealth of descriptions of observable social change and contributions to these changes, including Oxfam's and its partners' contributions. In addition to each project being able to draw contextual lessons learnt, Oxfam Novib's large-scale deductive analyses of qualitative data in all outcome statements enables assessment and validation of the theory of change. The results of the analysis in the Right to Food sub-programme will be presented.

**PN 16 Outcome Harvesting Customized for Monitoring and Evaluation in the Dutch Dialogue & Dissent Lobbying and Advocacy Programme for Development**

## **PC 065 - When Performance Evaluation Using DAC Criteria won't do: Findings from an Outcome Harvesting mid-term Evaluation of four NIMD Country Programs**

*N. Van der Jagt*<sup>1</sup>

<sup>1</sup> *The Netherlands Institute for Multiparty Democracy NIMD, Planning- Monitoring and Evaluation Coordinator, The Hague, Netherlands*

Nic Van der Jagt is the planning, monitoring and evaluation coordinator for The Netherlands Institute for Multiparty Democracy (NIMD). Nic will explain why performance evaluation using DAC criteria does not always work and the lessons learned from a pilot with the Outcome Harvesting alternative in the form of a mid-term evaluation of four NIMD D&D country programmes. NIMD is a democracy assistance organization that supports political parties in developing democracies through interparty dialogue whose effectiveness is especially difficult to measure with classical performance evaluations. He will present the experiences with an Outcome Harvesting in Mali and Mozambique, describing the extent to which the outcomes achieved represent patterns of progress towards respective programme objectives. Attention will be paid to the substantiation step of the OH methodology: How to determine if the outcome information is credible enough for learning and accountability.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 17 Great Expectations: An Enquiry into the Premise of Directing Institutional Investors to Emerging Markets, Especially in Times of Market Crises**

## Great Expectations: An Enquiry into the Premise of Directing Institutional Investors to Emerging Markets, Especially in Times of Market Crises

F. Korfker<sup>1</sup><sup>1</sup> *International Fund for Agricultural Development, Senior Advisor, Rome, Italy*

Introduction: To move the needle on emerging market development, private sector capital needs to be mobilized in billions through new investment vehicles with appropriate risk management and regulatory frameworks (Source: United Nations Financing for Development Conference, Addis Ababa, 2015). Global financial assets exceed \$300 trillion with nearly \$100 trillion invested with sovereign wealth funds and institutional investors such as pension funds, insurance companies and fund-of-funds (Source: Institute of International Finance, 2017). One of the most intriguing areas of development financing today is to find unique ways to channel the aforementioned trillions of commercial capital to developmental projects in emerging markets. In some instances, such long-term capital can potentially double up as an “accelerant” for a multilateral bank (MDB)’s or international financial institution (IFI)’s direct interventions to respond to a global or regional crisis. A recent strategy, pursued by MDB/IFIs in this context, is to position emerging market private equity both as an investable asset class that can generate positive returns for stakeholders and a contributor to development goals. The aim of this first presentation in the panel is three-fold: (a) to present methods and techniques to evaluate and benchmark the performance of funds, (b) to examine the potential of private equity funds as a crisis-response tool and (c) to share specific examples where such methods and instruments have been used to evaluate business lines within the World Bank Group. In addition, an ex-post case study will be presented by the second panelist, based on an evaluation conducted by EBRD entitled Crisis Response to Greek Bank Subsidiaries in Southeast Europe. This crisis response focused on recapitalizing the financial sector to provide liquidity to the private sector. The evaluation addresses cooperation across IFIs, the role of policy dialogue, and the relationships between borrowing financial institutions, parent financial institutions and the IFI providing crisis funding. The case study aims to provide lessons learned for use of funds and other vehicles in future crisis responses. Motivation: Raghavan Narayanan has many years of career experience in private sector, having recently pioneered a new benchmarking methodology to evaluate private equity investment funds. The case study by Barry Kolodkin aims to provide lessons from a recent and topical crisis to not only support the discussion of utilizing collective investment vehicles for future crises but to facilitate discussion on the applicability of evaluation techniques for the future. The third presenter will chair the session, summarize the key findings and takeaways, and facilitate the panelists interaction with the audience. The presentation will set out of the global context of institutional investors, the role of private equity instruments, application in emerging markets context, describe the various approaches to engage with private sector through the above instruments, detail the various methodologies used to analyze the interventions’ effects, with respect to the market crises environment and provide suggestions in a forward looking way. Presentations will be given along with case study examples, and it is expected that there will be a lot of interaction between the panelists and the participants.

**PN 17 Great Expectations: An Enquiry into the Premise of Directing Institutional Investors to Emerging Markets, Especially in Times of Market Crises**

## **PC 066 - Panel Contribution**

R. Narayanan<sup>1</sup>

<sup>1</sup> World Bank Group, Senior Evaluation Officer, Washington, USA

**PN 17 Great Expectations: An Enquiry into the Premise of Directing Institutional Investors to Emerging Markets, Especially in Times of Market Crises**

## **PC 067 - Panel Contribution**

B. Kolodkin<sup>1</sup>

<sup>1</sup> European Bank for Reconstruction and Development, Evaluations, London, United Kingdom

**PN 17 Great Expectations: An Enquiry into the Premise of Directing Institutional Investors to Emerging Markets, Especially in Times of Market Crises**

## **PC 068 - Panel Contribution**

F. Korfker<sup>1</sup>

<sup>1</sup> International Fund for Agricultural Development, Senior Advisor, Rome, Italy

Wednesday, 3 October 2018  
14:30–16:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 18 Evaluation Professionalisation in Practice: What are the Next Steps to Follow?**

## Evaluation Professionalisation in Practice: What are the Next Steps to Follow?

*T. Orfanidou<sup>1</sup>*

<sup>1</sup> Hellenic Association of Management Consulting Firms SESMA, Hellenic Association of Management Consulting Firms SESMA, Athens, Greece

Evaluation Professionalisation in Practice: What are the next steps to follow? A long discussion is in place, for several times, concerning the Evaluation Professionalisation. The subject has established a debating floor among evaluators, users, social stakeholders, and those setting the criteria for its design and implementation. According to EES and several experts in the field, including EES participants in the relative thematic working group, professionalism evokes expertise, credibility and concern for human welfare. The challenge involves accelerating the harmonization of ethical, quality and capability standards, increasing the autonomy and independence of evaluators and helping to ensure that evaluators have the qualifications needed to perform quality work. It also involves a well-defined body of knowledge, a set of specialized skills and a host of ethical guidelines. It displays trans-disciplinary features that allow it to support all the social sciences through a well-stocked tool kit of proven methods and processes. In line with the international and European guidelines, the GEA document, the outputs of several initiatives, workshops, panels and discussions evaluation professionalisation also recognizes the importance of

- (i) voluntariness;
- (ii) autonomy;
- (iii) legitimacy;
- (iv) pluralism;
- (v) transparency;
- (vi) equity; and
- (vii) quality assurance.

The idea to propose a panel discussion at the EES2018 biennial conference emerged during the discussions held on the occasion of a workshop organized by the Hellenic Evaluation Society in Thessaloniki (26/1/2018). The objective of the proposed panel is to identify the evaluation professionalisation achievements up to now, steps taken in practice, investigate the needs for action in the immediate future. The panel discussion will also aim at presenting a common framework agreed by the participants in the panel, facilitate a mutual understanding of the actions needed in the different aspects of evaluation professionalism. The panel will discuss actions for encouraging a quality culture and for identifying ways to elaborate efficient upgrade of skills, multidisciplinary use of methods and tools and exploitation of evaluation research and knowledge. The panel will contribute to the identification of the progress made in the field. It shall also discuss ways of improving capacity in the design, quality in the implementation and usefulness for social partners. The panel will try to address the following issues:

1. How do we evaluate the evaluations? What can be performed in practice?
2. How 'competent' are the commanders of the evaluation? How do we improve their capacity?
3. What are the next step in skills, competences and knowledge accreditation and enhancement?
4. What is the role of the consultants in this field?
5. How do we improve both implementation and use of evaluation?

The proposed panel will provide a sound opportunity for establishment of a constructive dialogue on evaluation professionalism, involving both decision makers and evaluation practitioners. It will explore possibilities in improving evaluation design and upgrading of skills and knowledge as a precondition for better implementing evaluation. It will also provide an opportunity for in-depth discussions on the need for establishing

- (i) voluntariness;
- (ii) autonomy;
- (iii) legitimacy;
- (iv) pluralism;
- (v) transparency;
- (vi) equity;
- (vii) quality assurance.

### PN 18 Evaluation Professionalisation in Practice: What are the Next Steps to Follow?

## PC 069 - The Case for Professional Designation in Evaluation

*L. Davies<sup>1</sup>*

<sup>1</sup> ICDC Inc, ICDC Inc, Ontario, Canada

The idea of “professionalising” evaluation practice in an attempt, inter alia, to contribute to its quality, has been evoked since at least the early years of some of the longer established evaluation societies and associations, e.g. the Canadian Evaluation Society in the late eighties.

That it took CES about 30 years to arrive at the point where it developed and implemented successfully its credentialing system, reflects not only the pioneering nature of the Canadian initiative, but as well the depth and breadth of the questions and considerations with respect to “professionalisation” that the discipline of evaluation continues to grapple with.

These issues are not unique to the debates among “evaluators”, they have been and continue to be, forged by the evolving history of interdependencies between the social exercise of remunerative practice, society’s demand for specialised knowledge and the political economy of “expertise”, its determination, control and market value.

A recurring question, both in the universe of the theory and practice of evaluation, as in the debate on its professionalisation, is that of the identity of the discipline, the “what is evaluation” koan that alternatively enlightens in moments of self-reflection and frustrates when attempting to explain it to a prospective client.

And it is this koan that is often the fulcrum of circular conversations among “evaluators”, where the cake and its eating are one and the same, conversations that keep the spin sufficient to maintain temporarily the illusion of stability, and of going nowhere fast.

The moment of inertia on the question of professionalisation has been a long one, although there are increasingly clear, and I would say encouraging, signs of wobble. Not because, in my view, of a reasoned and controlled development of the discussion, but rather because of reality’s usual and masterful way of making you pay before it teaches you the lesson. Or, in our case, reminds us of the lesson we failed to learn or to remember. “If you think a qualified professional evaluator is expensive try an unqualified one”.

If indeed evaluation can make a positive difference to society, it can also do harm. And this truth behoves us to nurture and protect the practice of evaluation so that it contributes positively to the sustainable betterment of humanity.

In “The logic of evaluation professionalism”, Bob Picciotto, identifies five “serviceable criteria for the assessment of the professionalism of a distinct occupational group.” Of these, two appear most important now: Expertise (high quality education, exposure to practice, theoretical knowledge, specialised skills, sound judgment, mastery of techniques). Credentials (degree from an accredited tertiary education establishment, professional designation, tested performance, membership in professional associations). As a number of other professions have done, evaluation should have a “common body of knowledge” consistent with the ever-evolving body of theory of evaluation as well as a common body of skills and abilities required of the professional evaluator. Evaluation should also employ a harmonized approach to testing stringently the appropriateness and sufficiency of professional evaluator knowledge, skills and abilities.

#### **PN 18 Evaluation Professionalisation in Practice: What are the Next Steps to Follow?**

### **PC 070 - Do Evaluators Need a Unique Set of Skills?**

*S. D'Errico*<sup>1</sup>

<sup>1</sup> *International Institute for Environment and Development, International Institute for Environment and Development, London, United Kingdom*

Do evaluators need a unique set of skills?

The objective of the proposed panel is to identify the evaluation professionalisation achievements up to now, steps taken in practice, and investigate the needs for action in the immediate future.

The panel discussion will also aim at presenting a common framework agreed by the participants in the panel, and facilitate a mutual understanding of the actions needed in the different aspects of evaluation professionalism.

The panel will also discuss actions for encouraging a quality culture and for identifying ways to elaborate efficient upgrade of skills, multidisciplinary use of methods and tools and exploitation of evaluation research and knowledge.

The panel will contribute to the identification of the progress made in the field. It shall also discuss ways of improving capacity in the design, quality in the implementation and usefulness for social partners.

The panel will try to address the following indicative issues:

1. How do we evaluate the evaluations? What can be performed in practice?
2. How ‘competent’ are the commanders of the evaluation? How do we improve their capacity?
3. What are the next step in skills, competences and knowledge accreditation and enhancement?
4. What is the role of the consultants in this field?
5. How do we improve both implementation and use of evaluation?

The proposed panel will provide a sound opportunity for establishment of a constructive dialogue on evaluation professionalism, involving both decision makers and evaluation practitioners. It will explore possibilities in improving evaluation design and upgrading of skills and knowledge as a precondition for better implementing evaluation. It will also provide an opportunity for in-depth discussions on the need for establishing

- (i) voluntariness;
- (ii) autonomy;
- (iii) legitimacy;
- (iv) pluralism;
- (v) transparency;
- (vi) equity; and
- (vii) quality assurance.

**PN 18 Evaluation Professionalisation in Practice: What are the Next Steps to Follow?****PC 071 - Three Levels of Professionalization**

*H. Simons<sup>1</sup>*

<sup>1</sup> *University of Southampton, Southampton, United Kingdom*

For over 20 years now evaluation societies have been exploring different strategies for enhancing the professionalization of evaluation. These include the generation of principles, standards, ethical statements, capabilities – and, more recently, various forms of peer review. These are all in addition to the ongoing professionalization efforts of societies to organize high level annual evaluation conferences and seminars, training programmes and supportive materials and generate evaluation journals which publish methodological and theoretical articles on evaluation. Many of these initiatives are focused on helping the individual evaluator or teams of evaluators enhance their skills and practice. Some are directed more towards the broader aim of establishing a body of knowledge and corporate identity for the field of evaluation as a social, ethical and political practice. Yet there is still more to be done to embed the importance of such a field within the wider society, so those who may benefit from evaluation come to recognize and value our services.

The presentation will address these three level of evaluation professionalization and lead into a discussion of how in particular to address this third level.

**PN 18 Evaluation Professionalisation in Practice: What are the Next Steps to Follow?****PC 072 - Young Emerging Evaluators: The Role of Higher Education in Building Their Capacity**

*M. Koutoulakis<sup>1</sup>*

<sup>1</sup> *University of Piraeus, M. Karaoli & A. Dimitriou St., International & European Studies Dpt., Piraeus, Greece*

According to EES and several experts in the field, including EES participants in the relative thematic working group, evaluation professionalism evokes aspiration and expertise. It also involves a well-defined body of knowledge, a set of specialized skills and a host of ethical guidelines. It displays trans-disciplinary features that allow it to support all the social sciences through a well-stocked tool kit of proven methods and processes.

For that, and especially for young emerging evaluators, the role of Higher Education Institutions to form, empower and build capacity of young evaluators is important. Although usually the knowledge provided in Universities is disparate and 'silent', a related mapping reveals that several learning objects in Universities can form a learning platform for empowering evaluation practice.

In the present paper we will try to challenge some basic issues:

- How ready are Higher Education Institutions in undertaking an active role to build capacity and empower young evaluators?
- How can Higher Education Institutions aspire these emerging evaluators? And in which topics?
- Can they, and to what extent, contribute in evaluation research?
- Can knowledge and skills, in essential methodologies, theoretical approaches and applications in evaluation and performance measurement, be built in an interdisciplinary way in specific curricula?
- What is the role of Higher Education Institutions in promoting evaluation of policies and practices at national and regional level? And for that in building a related capacity for their students?
- Can they set an evaluation platform for both theoretical and applied working knowledge to reveal that evaluation aspects are hidden on a variety of learning objects already present in some curricula?

- Can they contribute to build an understanding of the role that economic analysis plays in evaluation, including cost-benefit and cost-effectiveness analysis, econometric applications or of many other approaches?
- Can they actively participate in the evaluation community as partners, sharing knowledge and research outputs, for the benefit of all parties involved in the evaluation?

Wednesday, 3 October 2018  
16:30–18:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

PN 19 Evaluation for a Caring Society. The Potential of Care Ethics for Evaluation

## Evaluation for a Caring Society. The Potential of Care Ethics for Evaluation

*T. Abma*<sup>1</sup>

<sup>1</sup> VU University Medical Center, Medical Humanities- Amsterdam Public Health institute, Amsterdam, Netherlands

**Rationale:** Evaluation is a value-laden practice navigating in a complex world of various competing value-commitments. Ernest House (1993) was among the first who pointed out that in a capitalist society public organizations increasingly need to justify their practices through formal evaluations led by efficiency and economic productivity. Jennifer Greene and Helen Simons (2014) argued that evaluation may not fulfill its democratizing vision if it is not prepared to deliberately take into account a broad array of values, including social justice and equity. The concern that economic values may take privilege over other values led a group of evaluators and ethicists to explore the potential of care ethics for evaluation. This resulted in a volume in the series *Evaluation and Society*, entitled *Evaluation for a Caring Society*, edited by Merel Visse and Tineke Abma.

**Objectives sought:** Evaluators reflexively navigate in a complex web of values and value-commitments. The objective of this panel is to heighten the awareness of evaluators' responsibility to critically explore values embedded in evaluation practices and contexts and to present lessons drawn from our exploration of what it means to practice evaluation from an ethos of care.

**Brief narrative and justification (with regards to the review criteria above):** In our current neo-liberal context there is a broader concern that economic values may uncritically take privilege over other values. Zygmunt Bauman put it aptly: "A consumerist attitude may lubricate the wheels of the economy, but it sprinkles sand into the bearings of morality." (Bauman and Donskis, 2013, p. 150). In healthcare practices, for instance, care is largely evaluated in terms of a 'market' with 'consumers' that need to be satisfied with certain 'products'. Such evaluations have little to do with what it means to be a good doctor or what it means to provide good care for vulnerable patients and establish trust. Among professionals this may lead to alienation from moral responsibilities (Schwandt, 2002).

Tineke Abma will chair the panel, and present care ethics as a relational approach to morality (versus care as a skill, virtue or principle, or 'care work' like nursing, cooking). Care ethics is a way to navigate relationally, responsively, democratically and dialogically in particular practices to promote moral learning on care responsibilities. This is in line with Joan Tronto's (2012) recent plea to redefine democracy as developing shared understandings of care responsibilities; precisely, p. 30: "democratic politics should center upon assigning responsibilities for care, and for ensuring that democratic citizens are as capable as possible of participating in this assignment of responsibilities."

The panelists will show that care ethics offers a source of inspiration to enrich our evaluation practice. Jennifer Greene and Helen Simons will present their democratic, deliberative approach to evaluation and how they engage with a caring ethic in their relationships with people involved in the evaluation, like listening and being attentive to needs and vulnerabilities. Anders Hanberger will present a democratic 'caring' evaluation and reflect on 'practices of responsibility' in the context of program for refugee children.

**PN 19 Evaluation for a Caring Society. The Potential of Care Ethics for Evaluation**

**PC 073 - Democratic Caring Evaluation**

A. Hanberger<sup>1</sup>

<sup>1</sup> Umea University, Political Sciences, Umea, Sweden

The panellists will all present their vision on the potential of care ethics to enrich evaluation (each 10–15 minutes). Then we will start a discussion with the audience.

**PN 19 Evaluation for a Caring Society. The Potential of Care Ethics for Evaluation**

**PC 074 - Democratic Evaluation and Care Ethics**

H. Simons<sup>1</sup>

<sup>1</sup> University of Southampton, Education and Evaluation, London, United Kingdom

**PN 19 Evaluation for a Caring Society. The Potential of Care Ethics for Evaluation**

**PC 075 - Democratic Evaluation and Care Ethics**

J. Greene<sup>1</sup>

<sup>1</sup> University of Illinois, Education, Urbana-Champaign, USA

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 20 Evaluating Value for Money in Complex, Adaptive Development Programmes**

## Evaluating Value for Money in Complex, Adaptive Development Programmes

*J. King<sup>1</sup>, A. Hurrell<sup>2</sup>*

<sup>1</sup> *The University of Melbourne, Centre for Program Evaluation, Melbourne, Australia*

<sup>2</sup> *Oxford Policy Management Ltd, Monitoring & Evaluation, Oxford, United Kingdom*

As aid budgets come under increasing scrutiny from domestic and international stakeholders, the longstanding concern with developmental effectiveness has morphed into an urgent and commonplace obligation for donors and implementing partners to demonstrate Value for Money (VFM). Despite this, there remains a lack of appropriate methods to support meaningful VFM assessment in the international development sector, and a tendency to fall back on a collection of indicators of variable quality, devoid of an explicitly evaluative judgement. This panel presents an innovative new approach to VFM assessment which makes use of explicit evaluative reasoning – with rubrics setting out performance criteria (dimensions of VFM) and standards (levels of performance) tailored to the programme context – together with mixed methods (ranging from qualitative evidence to economic analysis) to support well-reasoned judgements about VFM that respond to donor accountability requirements. The approach balances the imperatives of efficiency and equity, ensuring the VFM assessment goes beyond financial indicators to consider impacts for marginalised groups, poor people, women and girls. In this way it positions VFM assessment to consider the impact of aid spending on gender equity and promoting more resilient societies. The approach builds on a theoretical framework proposed by King (2017) and operationalised in OPM's Guide to VFM assessment: King, J. (2017). *Using Economic Methods Evaluatively*. American Journal of Evaluation, March 2017. King, J. & OPM (2018). *OPM's Approach to Assessing Value for Money: A Guide*. Oxford: Oxford Policy Management Ltd. (<http://www.opml.co.uk/publications/opm's-approach-assessing-value-money>) The robustness of this intuitive and innovative approach to VFM analysis is demonstrated through an account of its application on the Sub-National Governance (SNG) Programme in Pakistan. A five year, £150m DFID-financed initiative, the SNG programme supports reforms in public financial management, planning, and service delivery improvement innovations, to support better democracy in Pakistan. It belongs to a category of adaptive and iterative governance reform programmes which are particularly poorly served by prevailing VFM methodologies, because their complexity, non-linear causal pathways, and expectations around being responsive to emerging opportunities and learning from successes and failures alike, cannot adequately be captured by quantitative indicators born of a log-frame. By demonstrating the how the approach was applied in this case, and how the emerging results were received and used, the panel will demonstrate the practicability and robustness of this evaluation-specific VFM methodology.

**PN 20 Evaluating Value for Money in Complex, Adaptive Development Programmes**

## PC 076 - Panel Contribution

*A. Hurrell<sup>1</sup>*

<sup>1</sup> *Oxford Policy Management Ltd, Monitoring & Evaluation Team Leader, Oxford, United Kingdom*

Alex will introduce the Pakistan Sub-National Governance (SNG) programme, to illustrate the challenges that arise when evaluating VFM in complex, adaptive programmes.

**PN 20 Evaluating Value for Money in Complex, Adaptive Development Programmes**

**PC 077 - Panel Contribution**

J. King<sup>1</sup>

<sup>1</sup> Director, Julian King & Associates Limited  
Honorary Fellow, The University of Melbourne  
Associate, Oxford Policy Management, Auckland, Australia

Julian will outline a novel approach to VFM assessment, developed and published in collaboration with OPM. He will illustrate its use in the SNG programme and show how it supports transparent judgements supported by sound reasoning and mixed methods evidence.

**PN 20 Evaluating Value for Money in Complex, Adaptive Development Programmes**

**PC 078 - Panel Contribution**

A. Hurrell<sup>1</sup>

<sup>1</sup> Oxford Policy Management Ltd, Monitoring & Evaluation Team Leader, Oxford, United Kingdom

Alex will explain the benefits of the VFM approach, in the SNG programme and more generally, as well as opportunities for further developing the method and its use.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 21 Evaluations to Transform: Addressing Gender Imbalances in Culturally Diverse Contexts**

## **Evaluations to Transform: Addressing Gender Imbalances in Culturally Diverse Contexts**

*F. Amariles<sup>1</sup>*

<sup>1</sup> *Learning for Impact, Chief Executive Officer, Weston, USA*

The Global Evaluation Agenda 2020 calls for every member of the evaluation community to play a role in attaining the EvalVision2020. This includes that VOPEs become more influential in enabling environment for evaluation within countries/regions with the overriding message of the SDGs “to leave no one behind”. Women and other Civil Society Organizations around the world have developed strategies and methodologies to mainstream gender in development, following directions from Beijing Platform (1995) and other international Agreements. But, unfortunately, these experiences have seldom been applied to evaluation.

Despite several advancements, the Gender Transformative approach does not get routinely settled in development evaluations. So, this panel promotes a pro-gender equality, culturally-responsive evaluation in the Global South, one that examines and questions power structures that can be – by action or omission – reinforcing and perpetuating inequalities.

Transformational issues in evaluation also require exploring new thinking about the role of evaluators as agents of social change. In turn, evaluators who wish to go beyond their traditional technical role to an advocacy-oriented and empowering role need to get involved and interact effectively with the culture of the communities where interventions being evaluated are operating. Furthermore, a gender transformative evaluation approach implies the development of some specific competences, strongly linked to evaluators' ways of being and ways of seeing, to effectively address and challenge traditional power structures and cultural norms that perpetuate gender gaps.

The main purpose of this panel is to enrich the process and to involve more stakeholders interested in strengthening a Gender Transformative approach to evaluation.

Four VOPEs (Voluntary Organizations for Professional Evaluation) representing three regions from the Global South (Africa, South Asia and Latin America) developed a short project to produce a curriculum and guidelines for mainstreaming a Gender Transformative approach into evaluation through capacity development (CD), based on existing best practices in each region and the development of technical, ethical, political and leadership competencies for application of the approach.

In each region, similar processes and tools to gather information were developed and adjusted. Each of the presentations will focus on regional processes and results, describing the path within and across regions and analyzing the process and lessons learned regionally as well as the collective results obtained.

**PN 21 Evaluations to Transform: Addressing Gender Imbalances in Culturally Diverse Contexts****PC 079 - South Asia Case Study Development for Gender Transformative Approach to Evaluation**

*S. Zaveri*<sup>1</sup>

<sup>1</sup> *Community of Evaluators South Asia, coordinator of GENSA, New Delhi, India*

Community of Evaluators South Asia contacted gender experts from various countries in South Asia to express interest in developing a gendered curriculum that is contextual. Capacity building needs were assessed to inform gender and evaluation training needs. Competencies that respond to South Asian realities were identified through a collegial and collaborative process. Based on geography with a special thrust to involve young evaluators, several ideas for the gender transformative curriculum were developed related to each country's context – Afghanistan, Nepal and India. This has resulted in identifying themes such as Gender Frameworks that particularly work in South Asia; Ethical principles; application of DAC criteria and their adaptation to gender and complexity in South Asia and barriers to gender participation in conflict countries. The South Asian experience leaned on the learnings from Latin America and Africa – both the regional associations having far more experience than CoE SA.

**PN 21 Evaluations to Transform: Addressing Gender Imbalances in Culturally Diverse Contexts****PC 081 - African Wide Experiences of Gender Transformative Approach in Evaluation**

*M. Jansen van Rensburg*<sup>1</sup>

<sup>1</sup> *AGDEN/Resilience Analysis Consulting, Independent Research Professional, Johannesburg, South Africa*

The Africa Gender and Development Evaluation Network (AGDEN) component for the South to South project includes an African wide investigation and search for cases of good practice. AGDEN follows a two-tiered process for the study. Two surveys are used to identify current capacity development and implementation of gender transformative evaluation. One is distributed to 35 VOPEs and the second individuals. These aim to identify priorities for capacity development. Additional interaction at the African Evaluation Association (AfrEA) and other international forums contributes to the information. This process enables the identification of case studies that address the curriculum needs that in turn will form the training presentation. The final step includes translation into French and English to ensure the implementation in both Anglophone and Francophone Africa.

**PN 21 Evaluations to Transform: Addressing Gender Imbalances in Culturally Diverse Contexts****PC 082 - Evaluations to Transform: What Competencies Do We Need to Address Gender Power Issues in Diverse Cultural Contexts? Lessons from Latin America**

*S. Salinas*<sup>1</sup>

<sup>1</sup> *Independent Evaluator, Coordinator RELAC, La Paz, Plurinational State of Bolivia*

Latin America is one of the most culturally diverse regions: indigenous exclusion and discrimination have recently been made visible and public responses have emerged to address its effects, while indigenous worldviews inspire new development paradigms, (i.e., the "Good Living" philosophy). Emerging counter-hegemonic worldviews, however, not necessarily tackle underlying power structures such as gender power relations. Transformational issues in evaluation require exploring new thinking about the role of evaluators as agents of social change. In turn,

evaluators who wish to go beyond their traditional technical. role to an advocacy-oriented and empowering role need to get involved and interact effectively with the culture of the communities where interventions being evaluated are operating. Furthermore, a gender transformative evaluation approach implies the development of some specific competences, strongly linked to evaluators' ways of being and ways of seeing, to effectively address and challenge traditional power structures and cultural norms that perpetuate gender gaps.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 22 Spreading the good news: applications of theory-based causal analysis in international development**

## Spreading the Good News: Applications of Theory-Based Causal Analysis in International Development

*E. Raimondo<sup>1</sup>, E. Stern<sup>2</sup>, J. Vaessen<sup>1</sup>, J. Schmidt<sup>3</sup>*<sup>1</sup> World Bank Group, Independent Evaluation Group, Washington, USA<sup>2</sup> Lancaster University, Educational Research, Lancaster, United Kingdom<sup>3</sup> German Institute for Development Evaluation, Methodology, Bonn, Germany

"There is nothing as practical as a good theory". This phrase, purportedly coined by Kurt Lewin in the 1940s and later taken up *inter alia* by Carol Weiss and Ray Pawson resonates well with most evaluators and development practitioners nowadays. Theories of change are omnipresent in International Development. Despite the ubiquity of program theories, there are continued misunderstandings and misconceptions about the use of theory in evaluation and planning. Some of the symptoms of this latent issue are: the mono-theory bias (a tendency to ask about "THE" theory), the intervention-centric bias (a tendency to relay external factors to the far periphery of the evaluation), limited questioning about the sources of theory, let alone the use of "big T" theories stemming from social sciences. Moreover, despite the groundbreaking 2012 DFID report authored by a team led by Prof. Stern on expanding the range of impact evaluation methods in international development, the potential of theory-based causal analysis remains underutilized in the field. Yet, some interesting real-world examples have emerged. To demonstrate the potential of theory-based causal analysis, this panel will elucidate four points about the practical relevance of this approach in international development. First, the panelists will show that a theory-based approach to causal analysis is truly "practical", in the sense that it offers depth of knowledge and understanding on the conditions under which a program may or may not be successful, generating information that are otherwise lacking in evaluative analysis. Second, the panelists' experiences will show that theory-based causal analysis does not need to be expensive. In fact, a variety of approaches can be adapted to different evaluation contexts and demands. Third, while theory-based causal analysis is not-method specific – it can be embedded in a variety of quantitative and qualitative approaches – the application of qualitative or quali-quantitative theory-based methods is especially underutilized. The panelists will show the value of these approaches. Finally, given the significant limitations to the applicability of (quasi) experimental approaches to causal analysis particularly with regards to the nature of the evaluand, the panels will show how theory-based approaches can fill an empirical gap. Structure of the panel: After an introduction by the chair, the three panelists will provide three lightning talks. The first will cover the main problems with the use of theory in international development; the second and third will provide two real-world examples of application of theory-based causal analysis, respectively discussing process-tracing and pattern-matching with and without Qualitative Comparative Analysis. The panel will then engage in a dialogue with the audience, moderated by the chair on broader considerations for future use of theory-based causal analysis in international development.

**PN 22 Spreading the good news: applications of theory-based causal analysis in international development****PC 083 - Three Case-Based Approaches in One: Combining Pattern Matching, Process Tracing and QCA Under Real-World Evaluation Conditions**

*E. Stern<sup>1</sup>, E. Raimondo<sup>2</sup>, J. Vaessen<sup>2</sup>, J. Schmidt<sup>3</sup>*

<sup>1</sup> Lancaster University, Educational Research, Lancaster, United Kingdom

<sup>2</sup> World Bank Group, Independent Evaluation Group, Washington, USA

<sup>3</sup> German Institute for Development Evaluation, Methodology, Bonn, Germany

Studying complex causal pathways often requires combining multiple approaches. In the evaluation of the World Bank's support to Carbon Finance we combined three theory-based and case-based approaches to fully study and test the theory of change. A combination of process-tracing and QCA, within the broader approach of pattern matching allowed us to study in minute details the causal mechanisms underpinning the intervention. The sampling strategy guiding the selection of the cases allowed us to reach modest generalizability. The presentation will touch upon the technical dimension of the design, as well as the practical implications of "making it happen" within the context of a real-world evaluation with time, budget, and skills constraints.

**PN 22 Spreading the good news: applications of theory-based causal analysis in international development****PC 084 - How to Increase Internal Validity in Theory-Based Evaluation? Integrating Comparative Case Studies and Theory-Testing Process Tracing**

*E. Raimondo<sup>1</sup>, E. Stern<sup>2</sup>, J. Vaessen<sup>1</sup>, J. Schmidt<sup>3</sup>*

<sup>1</sup> World Bank Group, Independent Evaluation Group, Washington, USA

<sup>2</sup> Lancaster University, Educational Research, Lancaster, United Kingdom

<sup>3</sup> German Institute for Development Evaluation, Methodology, Bonn, Germany

Theory-based evaluations (TBE) typically use comprehensive programme logics to assess the effectiveness of public policy interventions. Conventional TBE approaches enable us to evaluate complex programmes but struggle to attribute the observed effects to the particular intervention. Despite increasing discussions about how to increase internal validity of TBE for theory-testing purposes, there are only a few practical applications and the 'methodological bridge' is still under construction. This paper seeks to contribute to the ongoing methodological debate by presenting the integrated TBE-approach we applied in a recent development evaluation. The evaluation focuses on the exit from General Budget Support (GBS) and asks whether and how the exit affected former outcomes of GBS (i.e. public expenditure, public financial management, and domestic accountability) in the recipient countries. The applied approach combines a comparative case study design (four country cases) with in-depth case analysis using Theory-Testing Process Tracing (TTPT) in Malawi and Zambia. We describe how we conducted TTPT as suggested by Beach and Pedersen (2013) in four-steps: (i) develop hypothetical causal mechanisms, (ii) operationalize the causal mechanisms, (iii) collect evidence, and (iv) Bayesian updating. We demonstrate the added value of integrating TTPT into complex development evaluations and argue that adding in-depth analysis of key causal mechanisms in a sub-set of case studies helped increase internal validity at relatively low cost. Beyond the use for the specific evaluation presented in the paper, Process-Tracing can be a viable method for theory-testing in complex evaluations more generally.

**PN 22 Spreading the good news: applications of theory-based causal analysis in international development**

## **PC 085 - Panel Contribution**

*J. Vaessen*<sup>1</sup>

<sup>1</sup> *Independent Evaluation Group, Washington, DC, USA*

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 23 Conducting Impact Evaluation in Difficult Contexts Using Geo-spatial Analysis**

## Conducting Impact Evaluation in Difficult Contexts Using Geo-spatial Analysis

*H. Khaira<sup>1</sup>, D. Gilbreath<sup>2</sup>*

<sup>1</sup> IFAD, Independent Office of Evaluation, Rome, Italy

<sup>2</sup> Caucasian Research Resource Centers, Research, Tbilisi, Georgia

Conducting impact evaluations in changing and complex contexts can be a difficult task. The challenges to be confronted may be both external and internal. From having a strong logistics dependency on the field team to having difficulty recruiting qualified evaluators or data enumerators willing to travel and work in dangerous places. External limitations can include inaccessibility to sites and informants due to security concerns, reduced time in the field due to high resource costs, and collecting evidence that may not be of the same quality that had been originally envisioned. There is increasing recognition of the role that new information technologies can play in overcoming some of these challenges related to data collection in difficult contexts. This is made possible through more and better sources of remote sensing technologies, geo-referenced surveys, access to increased computing power and econometric techniques. Further, when used in combination, the positive features of these techniques can be actively harnessed to produce a rigorous, reinforced impact evaluation at a substantially lower time and financial cost. In this presentation we demonstrate how remote sensing was used on irrigation infrastructure several years following a conflict in the Republic of Georgia. We share the results of an impact evaluation that used a combination of quasi-experimental techniques with remote sensing data collection using satellite imagery. This is followed by a discussion of the methodological techniques that were used. The remote sensing was used both in conflict and non-conflict areas; the validity of the geo-spatial analysis can be gauged from the fact that its results in non-conflict areas were comparable with results of a household survey also administered in the same non-conflict areas. These results bolster confidence in remote sensing techniques and open up exciting opportunities to use them. Besides being used in fragile situations, such satellite-based techniques can also be used to cover remote places. Further, they can also be used to respond to the increasing call to estimate the impact of humanitarian assistance where the humanitarian assistance community has long asked for better evidence on how each dollar should be effectively spent.

**PN 23 Conducting Impact Evaluation in Difficult Contexts Using Geo-spatial Analysis**

## PC 086 - Conducting Impact Evaluation in Difficult Contexts Using Geo-spatial Analysis

*D. Gilbreath<sup>1</sup>*

<sup>1</sup> Caucasian Research Resource Centers, Research, Tbilisi, Georgia

Conducting impact evaluations in changing and complex contexts can be a difficult task. The challenges to be confronted may be both external and internal. From having a strong logistics dependency on the field team to having difficulty recruiting qualified evaluators or data enumerators willing to travel and work in dangerous places. External limitations can include inaccessibility to sites and informants due to security concerns, reduced time in the field due to high resource costs, and collecting evidence that may not be of the same quality that had been originally envisioned. There is increasing recognition of the role that new information technologies can play in overcoming some of these challenges related to data collection in difficult contexts. This is made possible through more and better sources of remote sensing technologies, geo-referenced surveys, access to increased computing power and econometric techniques. Further, when used in combination, the positive features of these techniques can be actively

harnessed to produce a rigorous, reinforced impact evaluation at a substantially lower time and financial cost. In this presentation we demonstrate how remote sensing was used on irrigation infrastructure several years following a conflict in the Republic of Georgia. We share the results of an impact evaluation that used a combination of quasi-experimental techniques with remote sensing data collection using satellite imagery. This is followed by a discussion of the methodological techniques that were used. The remote sensing was used both in conflict and non-conflict areas; the validity of the geo-spatial analysis can be gauged from the fact that its results in non-conflict areas were comparable with results of a household survey also administered in the same non-conflict areas. These results bolster confidence in remote sensing techniques and open up exciting opportunities to use them. Besides being used in fragile situations, such satellite-based techniques can also be used to cover remote places. Further, they can also be used to respond to the increasing call to estimate the impact of humanitarian assistance where the humanitarian assistance community has long asked for better evidence on how each dollar should be effectively spent.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 24 Building European Evaluation Capacity: The Role of Young and Emerging Evaluators**

## Building European Evaluation Capacity: The Role of Young and Emerging Evaluators

*M. Branco<sup>1</sup>, B. Montrosse-Moorhead<sup>2</sup>*

<sup>1</sup> *Independent Consultant, EES board- EES TWG on emerging evaluators, Porto, Portugal*

<sup>2</sup> *University of Connecticut, American Evaluation Association- EvalYouth, New York, USA*

**Rationale:** This panel will bring awareness to the global, regional movements of Young and Emerging evaluators (YEE) and provide valuable insights into the situation of young professionals entering the evaluation field.

**Objectives:** The presentations represent a concerted effort to raise global awareness upon the issues that affect young and emerging evaluators (YEE) across the globe, and call attention to gaps in these initiatives at regional and national levels. Together, the different papers will look forward to exploring commonalities and differences, promoting a collective discussion on pros and cons of these kinds of activities; current needs that are not covered; lessons learned; priorities for the future and potential partnerships.

**Justification:** During the Year of Evaluation, in 2015, the discussions around evaluation capacities and capabilities intensified, setting the conditions for the launch of The Global Evaluation Agenda 2016–2020, which makes clear that the development of evaluation capacity should fully correspond with the needs of young and emerging evaluators. For young professionals, finding the guidance to create a successful career in evaluation is a major challenge. It is also needed a stimulus for engaging the community of policymakers and evaluators to use innovative approaches, strategies, and methodologies; to attract young people and to take advantage of their ideas and energies. Finally, it is crucial to promote more active inclusion of the next generation of evaluators in the debate and priorities set for evaluation in the context of the Sustainable Development Goals (SDGs). Each individual presentation of the panel will have an approximate duration of 10 to 15 min. Between half an hour and 45 min will be left for discussion among panellists.

**PN 24 Building European Evaluation Capacity: The Role of Young and Emerging Evaluators**

## PC 087 - Why the Future of Evaluation Rests on Improving What Novice Evaluators Know and Can Do

*B. Montrosse-Moorhead<sup>1</sup>, A. Rishko-Porcescu<sup>2</sup>, B. Gauthier<sup>3</sup>, W. Meyer<sup>4</sup>, B. Baruch<sup>5</sup>, W. Felcis<sup>6</sup>*

<sup>1</sup> *University of Connecticut, American Evaluation Association- EvalYouth, New York, USA*

<sup>2</sup> *Independent Consultant, EvalYouth- EvalYouth ECA, Kiev, Ukraine*

<sup>3</sup> *Independent Consultant, Canadian Evaluation Society, Gatineau- Québec, Canada*

<sup>4</sup> *Saarland University, Center for Evaluation CEval, Saarbrücken, Germany*

<sup>5</sup> *RAND Corporation, RAND Europe, London, United Kingdom*

<sup>6</sup> *Independent Consultant, Polish Evaluation Society- EES- IOCE, Warszawa, Poland*

Launched in November 2015, EvalYouth is a global network committed to cultivating evaluator capacity and strengthening the role of youth in evaluation. Highlighted in this talk are the enabling conditions that gave rise to EvalYouth, the multi-stakeholder process that fed into its' initial creation, its' goals and objectives, and progress in meeting these objectives. This talk concludes with a brief discussion of the connection between EvalYouth, the EvalAgenda2020, and other global networks to help EES members, both new and seasoned, understand why the success of this movement is so important for the future of Evaluation.

**PN 24 Building European Evaluation Capacity: The Role of Young and Emerging Evaluators**

**PC 088 - Creating YEE initiative in ECA region (Eastern Europe, Central Asia & South Caucasus): Opportunities and Challenges**

*A. Rishko-Porcescu*<sup>1</sup>

<sup>1</sup> *Independent Consultant, EvalYouth- EvalYouth ECA, Kiev, Ukraine*

The paper presents the opportunities and challenges, which YEE in ECA region and the YEE initiative, EvalYouth ECA faced. The network made the mapping of the formal and informal opportunities in evaluation capacity building, and exist and potential auditory; and it is in the process of developing partnerships in the region for YEE support at the beginning of the evaluation profession.

**PN 24 Building European Evaluation Capacity: The Role of Young and Emerging Evaluators**

**PC 089 - A Comprehensive Map of European Evaluation Capabilities**

*M. Branco*<sup>1</sup>, *W. Meyer*<sup>2</sup>, *B. Baruch*<sup>3</sup>, *W. Felcis*<sup>4</sup>

<sup>1</sup> *Independent Consultant, EES board- EES TWG on emerging evaluators, Porto, Portugal*

<sup>2</sup> *Saarland University, Center for Evaluation CEval, Saarbrücken, Germany*

<sup>3</sup> *RAND Corporation, RAND Europe, London, United Kingdom*

<sup>4</sup> *Independent Consultant, Polish Evaluation Society- EES- IOCE, Warszawa, Poland*

Based on a meta-analysis of several European studies focused on training and teaching of evaluation, along with primary data, the presentation will contribute to: a systematization of the differences across countries, sectors and methods; map the challenges and check for overlaps; identify major players and trends and draw recommendations for future activities and strategies that could be put in place.

**PN 24 Building European Evaluation Capacity: The Role of Young and Emerging Evaluators**

**PC 090 - Is Professionalization of Evaluation a Global Movement? Can It Be? Ought It Be?**

*B. Gauthier*<sup>1</sup>

<sup>1</sup> *Independent Consultant, Canadian Evaluation Society, Gatineau- Québec, Canada*

In the grand scheme of things, discussions of the professionalization of evaluation are relatively recent. Why have they surfaced when they have? Do these discussions mean the same thing everywhere? Do they stem from the same forces? Do they share common goals? The Professionalization Task Force of the IOCE has chosen to focus on empowering VOPEs to build the professionalization discussion with their members rather than trying to impose a one-size-fits-all model. Is this the right approach?

**PN 24 Building European Evaluation Capacity: The Role of Young and Emerging Evaluators**

**PC 091 - Discussant**

*T. Schwandt*<sup>1</sup>

<sup>1</sup> *University of Illinois, American Journal of Evaluation, Illinois, USA*

After each paper is presented, Professor Schwandt will offer commentary on the papers presented, and raise several important questions for the panellists and the field at large to consider related to evaluator training, evaluation capacity building, and the professionalization of evaluation.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 25 From Measuring to Managing for Results: Building in-country Monitoring & Evaluation (M&E) Capacities and Systems**

## **From Measuring to Managing for Results: Building in-country Monitoring & Evaluation (M&E) Capacities and Systems**

*L. Martin<sup>1</sup>, G. Perez<sup>2</sup>, E. de Villalobos<sup>1</sup>, R. Sayed Khan<sup>1</sup>, K. Schrader<sup>3</sup>, E. Sibanda<sup>4</sup>, R. Seiwald<sup>1</sup>*

<sup>1</sup> International Fund for Agricultural Development, Operational Programming and Effectiveness unit OPE, Rome, Italy

<sup>2</sup> Centers for Learning on Evaluation and Results, Centers for Learning on Evaluation and Results Latin America, Ciudad de México- CDMX, Mexico

<sup>3</sup> HELVETAS, Technical Advisory Services, Zürich, Switzerland

<sup>4</sup> Itad Ltd., Social Protection and Livelihoods, Brighton, United Kingdom

IFAD is recalibrating itself to cater to the changing nature of development and demand from countries. First, IFAD developed the Program for Rural M&E (PRIME) in partnership with the CLEAR centres, which attempts to fill the existing M&E gap in the rural development sector by focusing on strengthening country capacities, and by developing a global M&E and impact assessment training and certification programme. Second, IFAD developed the AG-Scans, implemented together with the Swiss INGO HELVETAS and the UK consultancy company Itad. Both initiatives address the institutional dimension of M&E systems in countries and conduct assessments to identify gaps and develop action plans. This will allow IFAD government counterparts to implement targeted improvements to their M&E systems, in order to better manage for results within this sector. This initiative builds on the CAP-Scan methodology and will be adapted to the rural sector, but embracing all the pillars of the CAP-Scan methodology.

In the global efforts towards achieving the Sustainable Development Goals, the development community and governments have agreed on over 230 indicators to track progress. Tracking is needed for informed decision making. First, robust tracking is essential in finding solutions to challenges that are dynamic, such as those caused by climate change; for example, climate resilient agriculture is a moving target as climate patterns continue to mutate. Second, tracking is essential in adapting global solutions to specific contexts: e.g. addressing root causes of fragility has local cultural elements). Third, without robust data, governments and development partners cannot assess the trade-offs of pursuing multiple goals: e.g. more aggressive growth requires more energy and water, and can endanger forests.

So when development agencies talk about building capacity to monitor the SDGs, they are in fact implying much more than bean-counting. It is about instilling a culture of results that enables governments and development partners to learn from project implementation, to make timely mid-course corrections, and to refine proposed solutions regularly, moving away from rigid blueprints. It is ultimately about connecting measurement with management to do development differently.

Lessons learned from implementation of the programmes will be fundamental in assisting other fields in recognising the added value of assessing in-country systems, and training and certification in M&E to develop similar schemes.

In this session, participants will learn about advantages of embarking on such a capacity building and global certification scheme from different perspectives, as well as discussing the benefits of having the assessment of in-country M&E systems specific to agriculture and rural development. The innovative public-good character of the initiative will build on current practices in evaluation by making learning material on M&E and impact assessment in agriculture and rural development available to a greater audience.

**PN 25 From Measuring to Managing for Results: Building in-country Monitoring & Evaluation (M&E) Capacities and Systems****PC 092 - Prime: How to Build and Implement a Global Training on M&E in Rural Sector**

G. Perez<sup>1</sup>

<sup>1</sup> Centers for Learning on Evaluation and Results, Latin America, Ciudad de México- CDMX, Mexico

The intervention will highlight CLEAR's experience in the creation and implementation of a training program on M&E focused on the rural development sector. Despite the importance of the M&E systems, there are still significant gaps in the agricultural and rural sectors, since M&E training programs are more broad-based cutting across. PRiME is a global training and certification program on M&E and impact assessment that strengthens M&E knowledge, skills, and practices in the rural development sector, to improve results measurement and data collection for informed decision-making (5–10 minutes).

**PN 25 From Measuring to Managing for Results: Building in-country Monitoring & Evaluation (M&E) Capacities and Systems****PC 093 - PRiME: Establishing a Global Certification Framework on M&E in Rural Development**

R. Sayed Khan<sup>1</sup>

<sup>1</sup> International Fund for Agricultural Development, Operational Programming and Effectiveness unit OPE, Rome, Italy

Raniya Sayed Khan will highlight IFAD's experience in creating and implementing PRiME, a training and certification programme on monitoring and evaluation for M&E officers from IFAD's government counterparts of global scale. In this session, participants will learn about advantages of embarking on such a capacity building and global certification scheme from different perspectives, as well as discussing the benefits of having the assessment of in-country M&E systems specific to agriculture and rural development. The innovative public-good character of the initiative will build on current practices in evaluation by making learning material on M&E and impact assessment in agriculture and rural development available to a greater audience. (5–10 minutes)

**PN 25 From Measuring to Managing for Results: Building in-country Monitoring & Evaluation (M&E) Capacities and Systems****PC 094 - Linking Knowledge Management and Assessments of Capacities of In-Country M&E Systems**

K. Schrader<sup>1</sup>

<sup>1</sup> HELVETAS, Technical Advisory Services, Zürich, Switzerland

Kai Schrader as the Learning and M&E specialist of the AG-Scans, during the panel he will point out the importance of feedback loops for learning from M&E and the AG-Scans. He will address the institutional dimension of M&E systems in countries and explain how the project will conduct assessments to identify gaps and develop action plans. (5–10 minutes)

**PN 25 From Measuring to Managing for Results: Building in-country Monitoring & Evaluation (M&E) Capacities and Systems****PC 095 - Results from Conducting In-Country M&E Capacity Assessments**

E. Sibanda<sup>1</sup>

<sup>1</sup> Itad Ltd., Social Protection and Livelihoods, Brighton, United Kingdom

Ethel Sibanda will provide insights from first experiences from conducting assessment of M&E capacities in ministries relevant from the field. She will provide specific case studies and first results of the Ag-Scans. The overall goal is to increase the measurability of the SDGs related to agriculture. The Ag-Scan diagnostics will allow IFAD government counterparts to implement targeted improvement to their M&E systems allowing them to better manage for results in the rural sector. This initiative builds on the CAP-Scan methodology and will be adopted to the rural sector but embracing all the pillars of the CAP-Scan methodology: Leadership, Accountability and Partnerships, Monitoring and Evaluation, Planning and Budgeting, and Statistics. The uniqueness of the Ag-Scan initiative is its specificity to the rural sector and for agricultural development providing high potential of scaling-up opportunities across IFAD's portfolio. A participatory approach to assess capacities and elaborate action plans together with government counterparts will ensure ownership and strong commitment to implement them. These action plans will become an integral part of up-coming country strategies of IFAD wherein resources will be allocated to support their delivery, at least in part. The action plans will also catalyse international efforts to enhance country M&E capacities also by facilitating knowledge-sharing in international fora. (5–10 minutes)

**PN 25 From Measuring to Managing for Results: Building in-country Monitoring & Evaluation (M&E) Capacities and Systems****PC 096 - Linking IFAD's Results Agenda with Strengthened In-Country M&E Capacities**

E. de Villalobos<sup>1</sup>

<sup>1</sup> International Fund for Agricultural Development, Operational Programming and Effectiveness Unit OPE, Rome, Italy

Eloisa de Villalobos as a result specialist also working on corporate commitments of the results agenda will link our in-country support with corporate priorities. Without robust data, governments and development partners cannot assess the trade-offs of pursuing multiple goals: e.g. more aggressive growth requires more energy and water, and can endanger forests. So when development agencies talk about building capacity to monitor the SDGs, they are in fact implying much more than bean-counting. It is about instilling a culture of results that enables governments and development partners to learn from project implementation, to make timely mid-course corrections, and to refine proposed solutions regularly, moving away from rigid blueprints. It is ultimately about connecting measurement with management to do development differently. Eloisa will comment on both initiatives (PRiME and Ag-Scans) addressing the institutional dimension of M&E systems in countries and conduct assessments to identify gaps and develop action plans. This will allow IFAD government counterparts to implement targeted improvements to their M&E systems, in order to better manage for results within this sector. (5–10 minutes)

**PN 25 From Measuring to Managing for Results: Building in-country Monitoring & Evaluation (M&E) Capacities and Systems**

**PC 097 - IFAD's Approach to Strengthen Results-Based Management Systems in Partner Countries**

*R. Seiwald<sup>1</sup>*

<sup>1</sup> *International Fund for Agricultural Development, Operational Programming and Effectiveness unit OPE, Rome, Italy*

IFAD is working on different levels to improve the capacity of partner countries – all initiatives are reinforcing each other and serve one common objective. IFAD committed in its Eleventh Replenishment (IFAD-11) to embrace a culture of results across the organization and in its Member States, which will help transform resources into development results and providing improved accountability to taxpayers and citizens. Key activities to achieve this goal are to strengthen the focus on results, enhance self-evaluation and accountability for results and to improve M&E capacity in the rural sector. The SDGs require common efforts embracing a change of culture towards accountability for the delivery of results. Raphael will comment on the development of new and upcoming initiatives to strengthen in-country capacity systems on leadership and transparency within the field of rural development. (5–10 minutes)

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 26 What Makes A Good Policy? WFP Lessons on Policy Formulation and Practicalities**

## What Makes a Good Policy? WFP Lessons on Policy Formulation and Practicalities

A. Cook<sup>1</sup>

<sup>1</sup> World Food Programme, Office of Evaluation, Rome, Italy

Policies represent guiding principles to set directions within the organization. They highlight broad guidelines on actions that will achieve the organization's goals. They act as a course of action to guide and influence decisions that deal with day-to-day operational matters.

In the context of Sustainable Development Goals and implementation of Agenda 2030, evaluation plays a key role in supporting major development priorities by informing institutional and policy actions. Evaluations inform both policy/agenda-setting as well as operational implementation and daily decision-making.

The WFP Office of Evaluation (OEV) introduced Policy Evaluations (PE) as a new type of evaluation in 2008 to identify policy gaps or the need to update existing policies, as well as to provide evidence to positively influence policy-making and increase accountability for policy results. Since then, ten policy evaluations have been completed and a Top 10 Lessons for Policy Quality in WFP has been developed after a careful analysis of these reports. Although policy standards and practice are not defined in WFP, the lessons present 'good practice' on policy formulation and will be of practical use when drafting new WFP policies.

The lessons have been divided into two areas:

- Policy formulation, deals with the elements required to ensure a high-quality policy design
- Policy practicality, deals with the elements required to increase the likelihood of successful implementation.

The 13<sup>th</sup> European Evaluation Society Biennial Conference represents an opportunity to present the results of the analysis and the lessons learned and share with participants key elements of good practice on policy formulation and practicalities (drawing on learning from WFP policy evaluations) to ensure a high-quality policy design and to increase the likelihood of successful implementation.

**PN 26 What Makes A Good Policy? WFP Lessons on Policy Formulation and Practicalities**

## PC 098 - Panel Contribution

A. Cook<sup>1</sup>

<sup>1</sup> World Food Programme, Office of Evaluation, Rome, Italy

The WFP Office of Evaluation (OEV) introduced Policy Evaluations (PE) as a new type of evaluation in 2008 to identify policy gaps or the need to update existing policies, as well as to provide evidence to positively influence policy-making and increase accountability for policy results. Since then, ten policy evaluations have been completed and a Top 10 Lessons for Policy Quality in WFP has been developed after a careful analysis of these reports. Although policy standards and practice are not defined in WFP, the lessons present 'good practice' on policy formulation and will be of practical use when drafting new WFP policies. The lessons have been divided into two areas:

- Policy formulation, deals with the elements required to ensure a high-quality policy design
- Policy practicality, deals with the elements required to increase the likelihood of successful implementation.

The 13<sup>th</sup> European Evaluation Society Biennial Conference represents an opportunity to present the results of the analysis and the lessons learned and share with participants key elements of good practice on policy formulation and practicalities (drawing on learning from WFP policy evaluations) to ensure a high-quality policy design and to increase the likelihood of successful implementation.

#### **PN 26 What Makes A Good Policy? WFP Lessons on Policy Formulation and Practicalities**

### **PC 099 - Panel Contribution**

R. Desole<sup>1</sup>

<sup>1</sup> World Food Programme, Office of Evaluation, Rome, Italy

**Ramona Desole (Evaluation Analyst, WFP)** will introduce the WFP policy framework and coverage norms for policy evaluations; provide an overview of the 10 policy evaluation and geographic coverage that have contributed to the Top 10 Lessons for Policy Quality; and present the approach and methodology for synthesising lessons from the initial set of policy evaluations. She will highlight the findings on policy formulation, practicalities, and summary and how they are being used in different ways in WFP.

#### **PN 26 What Makes A Good Policy? WFP Lessons on Policy Formulation and Practicalities**

### **PC 100 - Panel Contribution**

M. Guinot<sup>1</sup>

<sup>1</sup> United Nations Development Programme, Independent Evaluation Office, New York City, USA

Mar Guinot (UNDP Evaluation Analyst) will present the approach and methodology for synthesising lessons from the initial set of policy evaluations. She will provide an overview of the process followed to identify the ten lessons.

#### **PN 26 What Makes A Good Policy? WFP Lessons on Policy Formulation and Practicalities**

### **PC 101 - Panel Contribution**

G. Duffy<sup>1</sup>

<sup>1</sup> World Food Programme, Office of Evaluation, Rome, Italy

**Gaby Duffy (WFP Senior Evaluation Officer)** will share lessons from recent experience in the Office of Evaluation WFP in using the Top 10 Lessons for Policy Quality in the evaluation of WFP policies focused on protection and humanitarian principles and access

Thursday, 4 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 27 Why Evaluation Theory Should Be Used to Inform Evaluation Policies: Reflections on the Theory-Policy-Practice Connection**

## Why Evaluation Theory Should be Used to Inform Evaluation Policies: Reflections on the Theory-Policy-Practice Connection

*S. Lemire<sup>1</sup>, C. Christie<sup>1</sup>, N. Stame<sup>2</sup>, M. Mark<sup>3</sup>, M. Marra<sup>4</sup>*

<sup>1</sup> University of California- Los Angeles, Department of Education and Information Studies, Los Angeles, USA

<sup>2</sup> University of Rome, La Sapienza, Rome, Italy

<sup>3</sup> Pennsylvania State University, Department of Psychology, State College, USA

<sup>4</sup> University of Salerno, Department of Public Policy, Naples, Italy

Many organizations now have evaluation policies. Because formal evaluation policies intend to frame evaluation practice, it is important to understand the aims of policies and their influence on practice. Evaluation theory is also intended to guide practice, therefore the extent to which policies are informed by theory is also important to consider, given the related purposes of evaluation theories and policies. Because evaluation policies have emerged relatively recently, the conceptual and empirical work on the theory-policy-practice connection is limited. The overall objective of the panel is to promote a stronger integration of evaluation theory in evaluation policies and, in effect, to strengthen the evaluation theory-policy-practice connection.

**PN 27 Why Evaluation Theory Should Be Used to Inform Evaluation Policies: Reflections on the Theory-Policy-Practice Connection**

## PC 102 - The Evaluation Theory-Policy-Practice Relationship

*S. Lemire<sup>1</sup>*

<sup>1</sup> University of California- Los Angeles, Department of Education and Information Studies, Los Angeles, USA

The first presentation will set the scene for the panel by offering a brief description of past and current exchanges on the theory-policy-practice nexus in evaluation. Informed by a forthcoming paper on this topic, the potential role and purpose of evaluation theory in the context of evaluation policies are considered. Finally, the presentation considers the structure and content of three illustrative evaluation policies of three major evaluation commissioners, emphasizing their stated purpose and scope, content and structure.

**PN 27 Why Evaluation Theory Should Be Used to Inform Evaluation Policies: Reflections on the Theory-Policy-Practice Connection**

## PC 103 - Practicing What We Preach – Making the Case for a Stronger Integration of Evaluation Theory in Evaluation Policies

*C. Christie<sup>1</sup>*

<sup>1</sup> University of California- Los Angeles, Department of Education and Information Studies, Los Angeles, USA

In the second presentation, the aim is to increase our attention to the intersection between evaluation theory and policy, by examining the use of evaluation theory in the evaluation policies of three major evaluation commissioners. Special attention is awarded methods, use, and valuing – three core dimensions of evaluation theory. The presentation concludes with a set of reflections on the different strategies for and benefits of a stronger integration of evaluation theory in evaluation policies.

**PN 27 Why Evaluation Theory Should Be Used to Inform Evaluation Policies: Reflections on the Theory-Policy-Practice Connection****PC 104 - Six Characterizations in Search of an Evaluation Theory-Policy- Practice Relationship**

M. Mark<sup>1</sup>

<sup>1</sup> Pennsylvania State University, Department of Psychology, State College, USA

I will briefly address several points, with only a mild allusion to Pirandello.

- (1) Evaluation practice can be based on evaluation theory, but in ways that do not make the evaluation theory obvious for practitioners. Thus, the evaluation theory could be influential, but not be consciously used by the practitioner.
- (2) Should the aim be for evaluation policy to incorporate evaluation theory? Or should evaluation theory aid practitioners' judgment as they move from (general) policy to (specific) practice?
- (3) If evaluation policy is to embody evaluation theory, which evaluation theory?
- (4) One character missing in the evaluation theory-policy practice relationship is training. At present, the need for evaluation practice exceeds the structures for and avenues to high quality training that includes evaluation theory. This situation leads to a weak theory-practice relationship.
- (5) Another needed character is research on evaluation, including further research on linkages across theory and policy and practice.
- (6) Finally, the "author", that is, the body that has the power to set evaluation policy, is also an important character in this theater of evaluation.

**PN 27 Why Evaluation Theory Should Be Used to Inform Evaluation Policies: Reflections on the Theory-Policy-Practice Connection****PC 105 - A Behavioral Design for Evaluation Policy Reform in Italy**

M. Marra<sup>1</sup>

<sup>1</sup> University of Salerno, Department of Public Policy, Naples, Italy

I will explore how the evaluation perspective (orientation to accountability and/or learning), methods and use are reflected in the policy choices concretely pursued within Italy's executive and legislative branches. The assumption is that the top-down administrative law tradition and the stringent constraints on public finance have deeply influenced the way central and local agencies as well as the parliament have designed, requested, and used evaluation services. Drawing on Behavioral Public Administration, I then put forward possible nudge strategies to promote bottom-up cooperation between managers and evaluators to foster staff's merit and orientation towards results; professionalization and dialogue among evaluators for improving measurements and indicators; and cross-agency coordination to strengthen government-wide evaluation initiatives. A behavioral policy reform design is particularly relevant for countries of civil law tradition – like Italy – with top-down performance regimes, which are resisted and perceived as red tape negatively impacting on performance.

**PN 27 Why Evaluation Theory Should Be Used to Inform Evaluation Policies: Reflections on the Theory-Policy-Practice Connection****PC 106 - Discussant**

N. Stame<sup>1</sup>

<sup>1</sup> University of Rome, La Sapienza, Rome, Italy

Nicoletta Stame will in her capacity as discussant for the panel summarize and consider the implications of the panel contributions for the profession and practice of evaluation.

Thursday, 4 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 28 How Much Rigor Is Enough? Impact Assessment For Better Decision Making And Increased Resilience**

## How Much Rigor is Enough? Impact Assessment for Better Decision Making and Increased Resilience

*M. Branco<sup>1</sup>, A. Richards<sup>2</sup>*

<sup>1</sup> *www.mariana-branco.com, EES Board- EES Emerging Evaluators- EES Social Media, Porto, Portugal*

<sup>2</sup> *Social Value UK, Research and International Training, Liverpool, United Kingdom*

**Rationale:** This panel seeks to trigger a discussion around the contribution of social impact assessment to adaptive management and social resilience. Moreover, it aims to promote a debate amongst a variety of impact analysts coming from different schools of thought (e.g.: experimental studies, financial analysis, participatory evaluation) and backgrounds (e.g.: impact investment, venture philanthropy, NGOs, private sector, public sector) about the level of rigour necessary in impact evaluation for effective decision making in times of humanitarian, financial, political and social crises.

**Objectives:** Compare the different perspectives of rigor across sectors and methods against 3 criteria: relevance, completeness and accuracy; Understand how different impact assessment methodologies and levels of rigor can contribute (or not) to better decision making; Discuss how to choose amongst different impact evaluation approaches to identify the best opportunities to strengthen people's resilience and how these methods can be most useful in times of financial, environmental, political and social crises; Enhance the debate on quality standards in social impact assessment.

**Justification:** The world is constantly changing and we all need to be adaptable if we want to keep up. Impact measurement is a valuable tool to support the overall effectiveness of the intervention and to ensuring that as much is done with the resources available, for the benefit of society. However, for us to change and create more resilient societies, impact measurement needs to be responsive. The question is not whether or not to measure impact, but how to do it effectively and efficiently, and how to ensure that it is embedded into the organisation. It is crucial to ensure that impact measurement is done in a practical way. The primary motivation for measuring impact should not be to prove whether or not you're impactful for donors – though that's important too – rather, it's about using data to continually learn about the intervention and work to improve its impact. Only this can foster a decision-making process that strengthens resilience. However, for many, practicality is often seen as an antonym of "rigour" although rigour does not mean the same for every person. There are those who see rigour as statistical significance and control groups; the ones to whom rigour is in substantive financial and tangible economic data; and those who privilege the rigour of meanings and stakeholders narratives. This panel will explore different impact measurement methods and case studies, considering their level of rigour and how they contributed to significant decisions on resource allocation. To help structure the discussion three categories of rigour (relevance, completeness and accuracy) in impact measurement will be compared against their practical outcomes. Each speaker will present for 10 minutes. Individual presentations will be followed by a 30 minutes discussion.

**PN 28 How Much Rigor Is Enough? Impact Assessment For Better Decision Making And Increased Resilience**

**PC 107 - Comparative Analysis of Different Levels of Rigor when Measuring Social Return on Investment (SROI)**

*M. Branco<sup>1</sup>, A. Richards<sup>2</sup>*

<sup>1</sup> *www.mariana-branco.com, EES Board- EES Emerging Evaluators- EES Social Media, Porto, Portugal*

<sup>2</sup> *Social Value UK, Research and International Training, Liverpool, United Kingdom*

Compare social return on investment (SROI) against the different 3 criteria: relevance, completeness and accuracy; Provide examples of different levels of rigor in SROI and how they contributed (or not) to decision making; Discuss how SROI can contribute to promote social resilience and be useful in times of financial, environmental, political and social crises.

**PN 28 How Much Rigor Is Enough? Impact Assessment For Better Decision Making And Increased Resilience**

**PC 108 - Broadening Boundaries to Improve Estimates of Program Impact**

*M. Bamberger<sup>1</sup>*

<sup>1</sup> *The Word Bank, Evaluation, Michigan, USA*

The multiple dimensions of boundaries; How can broadening evaluation boundaries can increase the social rate of return; The challenge of boundaries in multidisciplinary evaluations; Boundaries and resilience; An example: applying boundaries in the evaluation of gender outcomes.

**PN 28 How Much Rigor Is Enough? Impact Assessment For Better Decision Making And Increased Resilience**

**PC 109 - Rigour Used by CBA in the EU Cohesion Policy Context and Stories About How is CBA Influencing (or not) Decision-Making**

*C. Pancotti<sup>1</sup>*

<sup>1</sup> *Development and Evaluation Unit- Centre For Industrial Studies CSIL and University of Milan, Member of The Society for Benefit-Cost Analysis, Milan, Italy*

CBA in the EU cohesion policy context: how a set of commonly shared working rules improved the consistency and rigour in cost benefit analyses for ERDF and Cohesion Fund applications; How CBA influences decision making: evidence from a recent ex-post evaluation; Retrospective CBA as a means to improve quality and rigour of ex-ante CBA.

**PN 28 How Much Rigor Is Enough? Impact Assessment For Better Decision Making And Increased Resilience**

**PC 110 - Triple Bottom Line: Is the Perfect Enemy of the Good?**

*E. Sarmiento<sup>1</sup>*

<sup>1</sup> *Nova School of Business and Economics, Novafrica, Lisbon, Portugal*

- Sharing value: integrating business and social analytics
- Target setting and measurement standards
- Top-down and bottom-up approaches to metrics
- Connecting with SDGs
- Counterfactual Impact Evaluation

**PN 28 How Much Rigor Is Enough? Impact Assessment For Better Decision Making And Increased Resilience****PC 111 - Apples and Oranges? Thoughts on Approaches of Impact Investing Community and MDB Private Sector Operations***R. Narayanan<sup>1</sup>*<sup>1</sup> *The World Bank, Independent Evaluation Group, Washington DC, USA*

- Scale and scope
- Standards and frameworks
- Trends and ways forward

**PN 28 How Much Rigor Is Enough? Impact Assessment For Better Decision Making And Increased Resilience****PC 112 - The Adoption of Social Risks Framework to Support Social Return on Investment Analysis: a Client Testimony***T. Culhari<sup>1</sup>, F. Serejo<sup>1</sup>, M. Branco<sup>2</sup>, H. Ricardo Lamas Diogo<sup>3</sup>, N. Gomes<sup>1</sup>, L. Ribeiro Queiroz de Araújo<sup>1</sup>*<sup>1</sup> *Social Corporate Responsibility, Voltalia Energia do Brasil, Natal, Brazil*<sup>2</sup> *www.mariana-branco.com, EES Board, Emerging Evaluators & Social Media, Porto, Portugal*<sup>3</sup> *Social Corporate Responsibility, Ramboll Environ, Salvador, Brazil*

Considering the social reality of increasing complexity and associated dynamics, integrative approaches have become an imperative in social impact assessment. In this presentation we will discuss the potential of combining Social Return on Investment (SROI) with risk assessment and management. SROI is an internationally recognized framework and methodological process to measure social impacts. Social risk management has been widely undertaken for business purposes, mainly to identify and manage risks related to stakeholders who can potentially affect a private or public initiative. Hence, the conventional social risk assessment typically puts more effort into assessing the consequences to the company rather than the impacts of the business activities to the communities. However, recent literature has been increasingly discussing the differences between social risks and business risks. For example, risk assessment frameworks have been used to structure human rights issues to address compliance against the United Nations Guiding Principles on Business and Human Rights. International Banks and donor trust funds are also amongst some of the leading agencies using risk analysis to account for their social investments. Nevertheless, what is not yet sufficiently explored is the use of risk analysis as a tool for adaptive management in the context of blended finance. During the implementation of a SROI analysis contracted by Voltalia Energia do Brazil, we identified the importance of combining both tools in order to increase social risk awareness and project responsiveness capacity. In this presentation the evaluation contractor together with the evaluation team we will:

- Discuss how can risk analysis be used to increase social businesses responsiveness to challenges;
- Detail our experience presenting practical tools to systematize and present the risks within a social impact assessment;
- Debate the importance of stakeholders' engagement in collecting social risk data;
- Describe how systematizing risks by investing in data visualization helps facilitate the communication between the SCR manager and the board of directors, enabling more effective decision making ;
- Explain how adopting change management measures has increased project ownership;
- Discuss how risk assessment contributes to mitigate negative impacts;
- Explore the potential of risk assessment to take into consideration local culture, needs, relationships and conflicts promoting better social resilience alongside with better planning and better project design.

- The evaluation team in conjunction will deliver the presentation of this paper with the client and project manager. We expect this to provide a more realistic experience to the audience and to contribute to advocate for the use of evaluation findings.

Thursday, 4 October 2018  
08:30–10:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 29 The Voluntary Evaluator Peer Review Model: Building Professional Accountability Through Professional Development**

## The Voluntary Evaluator Peer Review Model: Building Professional Accountability Through Professional Development

*W. Felcis*<sup>1</sup>

<sup>1</sup> Freelance evaluator, Riga, Latvia

**Rationale:** As a profession, evaluation generally still lacks robust structures for ongoing professional education and for evidencing practitioner competence. Currently, evaluation professional bodies are exploring ways in which these features of a profession can be provided to practitioners. Several evaluation associations and societies, including EES, have developed competencies or capabilities frameworks. The Canadian Evaluation Society (CES) has introduced a voluntary credentialing programme, and the AEA is currently exploring options for evaluator designation. An alternative, represented in the recently piloted EES Voluntary Evaluator Peer Review (VEPR) programme, focuses on identifying areas where practitioners need to build their skills and knowledge and canvassing ways to do so through a quasi-mentoring process involving a peer review. This panel discussion will explore two ways in which evaluation practitioners can assess their own competence – through auditing their level of ability against a structured set of capabilities, or through discussion of self-selected competencies with senior evaluation practitioners in a non-judgmental review process.

**Objectives sought:** The goals of the panel discussion are to present and assess the relative pros and cons of two conceptually different approaches to assessing one's professional capabilities as an evaluation practitioner, both involving self-reflection:

- (i) by auditing one's level of competence against each of the dimensions in the EES Capabilities Framework, and
- (ii) by taking part in a peer review process with selected senior practitioners.

**Brief narrative and justification:** Panellists will debate the advantages and problems with each approach outlined above. Audience comment will be invited, including ideas for alternative approaches to facilitating evaluation practitioner accountability for professional competence. Workshop participants will be engaged actively in two possible approaches to assessing one's professional evaluation capabilities, both approaches involving self-reflection. Through these exercises, participants will have an opportunity to consider the potential benefits, and possible drawbacks, of each approach. The following plenary discussion will focus on how these approaches, or others, may be used by VOPE's to encourage evaluation practitioners to firstly undertake regular, and rigorous, assessment of their competence in areas relevant to their evaluation work, and also develop a habit of personalised, ongoing professional development.

**PN 29 The Voluntary Evaluator Peer Review Model: Building Professional Accountability Through Professional Development**

## PC 113 - Pros and Cons of an Auditing Approach to Capabilities Assessment – Reviewee Perspective

*P. Oliver*<sup>1</sup>

<sup>1</sup> Pam Oliver Ltd Research and Evaluation, Director, Waiheke Island, New Zealand

5–6 mins.

**PN 29 The Voluntary Evaluator Peer Review Model: Building Professional Accountability Through Professional Development**

## **PC 114 - Pros and Cons of a Peer Review Approach to Capabilities Assessment – Reviewee Perspective**

M.A. Palenberg<sup>1</sup>

<sup>1</sup> Institute for Development Strategy, Director, Munich, Germany

5–6 minutes.

**PN 29 The Voluntary Evaluator Peer Review Model: Building Professional Accountability Through Professional Development**

## **PC 115 - Pros and Cons of Each Approach – Reviewer Perspective**

C. Rodriguez-Ariza<sup>1</sup>

<sup>1</sup> FAO, Programme Director, Addis Ababa, Ethiopia

5–6 minutes.

**PN 29 The Voluntary Evaluator Peer Review Model: Building Professional Accountability Through Professional Development**

## **PC 116 - Transportability of VEPR Across Cultures**

A. El Khoury de Paula<sup>1</sup>

<sup>1</sup> Independent evaluator/researcher, Madrid, Spain

5–6 minutes.

Thursday, 4 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### PN 30 New Evaluation Approaches for Changing Feminist Times

## New Evaluation Approaches for Changing Feminist Times

*P. Alvarez*<sup>1</sup>

<sup>1</sup> UN Women, UN Coordination Division, New York, USA

We are living in interesting times. A new awareness is emerging calling for a broad and deep societal change. Women are talking with a louder collective voice. Feminist lens make now possible that invisible agendas can be seen. This panel seeks to discuss, analyze and understand some of the implications of this new energy to consolidate evaluation agendas and methodologies that can benefit from it. As evaluators, we have to ask ourselves what we want to achieve and whether there are emerging new forms of speaking truth to power or whether they will be co-opted by hegemonic interests. This panel examines what new evaluation approaches can capitalize on this political movement to re-introduce feminist issues in the evaluation agenda. It will consider as well new methodologies, new rules and principles for evaluations in charged political times. What are the opportunities and the challenges? There is the diagnosis of an epidemic and the space to talk about it in ways that were not there before. For feminist evaluators and evaluators at large, the questions is whether there is a translation of this shift in evaluations and in the institutions in which or about which we produce evaluations. There are calls for new research agendas and new educational agendas. What are the changes that a new evaluation agenda needs to undertake in light of this societal change? What spaces are opening up for feminist evaluators? How much needed is feminist evaluation to ensure that evaluations are in synch with times? There seems to be no clear pathways for change. People are listening in different ways and a lot of questions are emerging: Is this a shift in perspective? A re-foundational moment in history? A tipping point? Another one? Will there be a backlash? Is this going too far? How cross-generational, cross-ethnicity, cross-class is this movement? How real is this? What are the calls for action? How this energy will be translated into long-lasting change? There are new opportunities to speak truth to power and to create a new normal. What can we do from the evaluation field to promote structural change? How can we learn from this explosion of energy, the amplification of the private into the public, the unstoppable determination for change? Feminist evaluation has been traditionally marginal and contested from the mainstream. An re-energized feminist movement is changing political agendas and media attention. Whether this is a long-lasting change or just another media moment, there is an opportunity to claim space for feminist issues. The evaluation world needs to remain in tune with political and public agendas to remain relevant.

### PN 30 New Evaluation Approaches for Changing Feminist Times

## PC 117 - Inclusive Systemic Evaluation for Gender equality, Environments and Marginalized voices (ISE4GEMs): A New Approach for the SDG Era

*S. Reddy*<sup>1</sup>

<sup>1</sup> UN Women, Independent Audit and Evaluation Service, New York, USA

With the advent of the United Nation's 2030 Sustainable Development Agenda, and its 17 Sustainable Development Goals, there is a call for more methodologies to understand and evaluate combinations of these global challenges, their integrated nature and their complexities. Accordingly, UN Women Independent Evaluation Office along with Australian and American researchers have written and piloted a new evaluation guide: *Inclusive Systemic Evaluation for Gender equality, Environments and Marginalized voices (ISE4GEMs): A new approach for the SDG era (2018)*. Referred to as the 'ISE4GEMs', this guide is an original piece of work that

brings together transdisciplinary evaluation methods, re-thinks systemic evaluation methodology and introduces the Gender equality, Environments and Marginalized voices (GEMs) framework. This contribution provides an amalgam of theoretical concepts that have been synthesized from systems thinking, social and ecological sciences to produce process guidance to address complexity and truly integrate transformative gender equality objectives into programming and evaluation work.

#### PN 30 New Evaluation Approaches for Changing Feminist Times

### PC 118 - Democratic Evaluation and Democracy: Exploring the Reality

*E. Lewis*<sup>1</sup>

<sup>1</sup> *University of Hull, Centre for Systems Studies, Beverley, United Kingdom*

The SDG Era requires evaluators to think systemically, systematically and intersectionally. A new paradigm is emerging that starts with the premise that each intervention is an opportunity for learning how to influence desired social change towards gender equality, sustainability, human rights and peace. The *Inclusive Systemic Evaluation for Gender equality, Environments and Marginalized voices (ISE4GEMs)* constitutes a new approach for the SDG era. ISE4GEMs responds to the challenges posed by the SDGs and the complexity of broad, intelinked agendas.

#### PN 30 New Evaluation Approaches for Changing Feminist Times

### PC 119 - New Evaluation Approaches for Changing Feminist Times: What Evaluation Agenda for the Fourth Feminist Wave?

*S. Reddy*<sup>1</sup>, *P. Alvarez*<sup>2</sup>, *D. Podems*<sup>3</sup>

<sup>1</sup> *UN Women, Independent Evaluation Service, New York, USA*

<sup>2</sup> *UN Women, Independent Evaluation Office, New York, USA*

<sup>3</sup> *OtherWISE Research and Evaluation, Evaluation, Cape Town, South Africa*

What some have called the fourth feminist wave is characterized by the use of social media and the “incredulity that certain attitudes can still exist”. This contribution wants to interrogate the feminist agenda for evaluations to better understand what is more helpful to keep momentum and capitalize on this enhanced attention to feminist issues.

For some, this new feminist energy is too focused on sexual harassment, toxic masculinities and gender-based violence. For others, this feminist wave is restricted to digital natives and urban elites. Feminist evaluations are both technical and political exercises. As such, feminist evaluations can contribute to build evidence on what works and consolidate the cultural shift in its analysis of public policy. Technically, feminist evaluations require more sophistication to be truly inclusive and address the complexity associated with broader social movements that seek to effect large scale change.

This panel contribution will explore the needs and opportunities for feminist evaluations to mobilize critical issues that can contribute to further its political scope.

Thursday, 4 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 31 Adapting Outcome Harvesting for Monitoring and Learning**

## **Adapting Outcome Harvesting for Monitoring and Learning (M&E)**

A. Tiernan<sup>1</sup>, R. Wilson-Grau<sup>2</sup>, J. Bravo-Hernandez<sup>3</sup>, H. Bach<sup>4</sup>, R.R. Ojok<sup>5</sup>

<sup>1</sup> Christian Aid Ireland, Programme Development Unit, Dublin, Ireland

<sup>2</sup> Ricardo Wilson-Grau Consultoria em Gestão Empresarial Ltda, Presidency, Rio de Janeiro, Brazil

<sup>3</sup> Red Cross Canada, Quality Assurance- Learning and Innovation, Ottawa, Canada

<sup>4</sup> ActionAid Denmark, Accountability Unit, Copenhagen, Denmark

<sup>5</sup> Tax Justice Network Africa, Institutional and Partnership Development, Nairobi, Kenya

Outcome Harvesting, as an alternative evaluation methodology to linear, predictive evaluation models, has been adapted by organisations around the world for M&E of what has been achieved and how interventions can learn about what is working in real time. In this session, Ricardo Wilson-Grau will introduce the methodology and its recent uses in the sector. Four colleagues will then present their experiences introducing and adapting Outcome Harvesting as an M&E tool in their organisations. Juliana Bravo is an advisor, evaluation, monitoring and learning with the Canadian Red Cross's International operations. Her presentation will showcase the Canadian Red Cross's application of Outcome Harvesting principles to support the customisation of the approaches six steps for monitoring and evaluating activities of a strategic partnership project aiming to improve capacity building for emergency response in five Red Cross National Societies in the Americas. This project is CERA, Capacity Building for emergency response in the Americas, a CRC-GAC Strategic Partnership project. Helene Bach is the monitoring and evaluation coordinator for ActionAid Denmark and responsible for integrating Outcome Harvesting AADK's monitoring and learning system. She will share how this experimental learning journey of AADK began with an external Outcome Harvesting evaluation followed by an internal learning review of AADK's youth engagement. The evaluation and review formed the basis for subsequently integrating Outcome Harvesting into the AADK M&E system, which entailed developing/usage of a database, adapting reporting systems as well as supporting staff and partners to regularly harvest outcomes. In parallel experience was shared with the broader CSO community in Denmark. Reagan Ronald Ojok is the monitoring, evaluation, accountability and learning officer with the Tax Justice Network Africa. He will explain the practicalities of harvesting outcomes in a workshop that doubles as a capacity-building exercise, and thus serves as an effective approach in determining the social change the TJNA network has influenced and the contribution of the participating institutions in Uganda; Tanzania; Zambia; Nigeria and Ghana. Alix Tiernan is the programme performance advisor at Christian Aid Ireland and part of the team that re-designed a seven country, Irish Aid funded governance and human rights programme, to take on an explicit focus on adaptive programme management. This re-design has involved developing a new approach to M&E, where Outcome Harvesting has become the main annual data collection methodology staff and for local partner organisations. Outcomes harvested are used for ongoing reflection about whether the programme strategies as set out in the Theory of Change are working. This process, called 'Strategy Testing', allows the programme to respond flexibly to changes in the context, but more importantly, it allows for critical reflection when a strategy does not seem to be producing the results expected and should therefore be adapted to be more likely to be effective, or stopped entirely. Christian Aid Ireland has developed an app to help use Outcome Harvesting for this type of monitoring.

**PN 31 Adapting Outcome Harvesting for Monitoring and Learning****PC 120 - Applications and use of Outcome Harvesting to Strengthen Monitoring, Evaluation and Learning for Canadian Red Cross projects**

J. Bravo-Hernandez<sup>1</sup>

<sup>1</sup> Red Cross Canada, Quality Assurance- Learning and Innovation, Ottawa, Canada

This presentation will showcase the Canadian Red Cross application of Outcome Harvesting principles supporting the monitoring and evaluation activities of an important strategic partnership project aiming to improve capacity building for emergency response in five Red Cross National Societies in the Americas.

**PN 31 Adapting Outcome Harvesting for Monitoring and Learning****PC 121 - The Outcome Harvesting Learning Journey of AADK – From Undertaking an External Evaluation to Building an OH Community in Denmark**

H. Bach<sup>1</sup>

<sup>1</sup> ActionAid Denmark, Accountability Unit, Copenhagen, Denmark

In the beginning of 2017 AADK commissioned an external Outcome Harvesting evaluation of our Tax Justice program. Also, we undertook an internal learning review of our youth engagement. Through highly participatory processes, outcomes were harvested in workshops and virtually from the change agents involved in the program (staff of ActionAid Denmark and ActionAid International, as well as ActionAid country offices, partners and youth trainers). The evaluation and review formed the basis for subsequently integrating Outcome Harvesting into the M&E system, which entailed developing/usage of a database, adapting reporting systems as well as supporting staff and partners to regularly harvest outcomes. In parallel experience was shared with the broader CSO community in Denmark, where AADK co-facilitated 4 trainings.

**PN 31 Adapting Outcome Harvesting for Monitoring and Learning****PC 122 - Harvesting Outcomes in a Workshop; Ideas and Practice from a Cross-Border Network of Civil Society Organisations**

R.R. Ojok<sup>1</sup>

<sup>1</sup> Tax Justice Network Africa, Institutional and Partnership Development, Nairobi, Kenya

The challenge of joint cross-border implementation of projects among network organisation has promoted the innovation and creativity, as realised through the introduction of Outcome Harvesting at network member level. This paper presents experience and reflection on the case of the Scaling Up Tax Justice (SCUT) project implemented by Tax Justice Network Africa (TJNA). The SCUT is being implemented in Uganda, Tanzania, Zambia, Nigeria and Ghana by five participating members of the network. The project seeks to build enhanced and sustained capacity of members to contribute to the achievement of the network's overall strategic mandate. This is being done through joint research, advocacy and training activities. Harvesting in a workshop is one of the most effective approaches in determining the social change, its merits, and the contribution of the participating institutions to the claimed outcomes. This approach also enhances the opportunity for capacity building as well as promotion of a learning culture among the partners.

**PN 31 Adapting Outcome Harvesting for Monitoring and Learning****PC 123 - Using Outcome Harvesting for Adaptive Management in Human Rights and Governance Programming**

*A. Tiernan*<sup>1</sup>

<sup>1</sup> *Christian Aid Ireland, Programme Development Unit, Dublin, Ireland*

Christian Aid Ireland has recently re-designed its Irish Aid funded governance and human rights programme to take on an explicit focus on adaptive programme management, informed by discourse around 'Doing Development Differently', 'Thinking and Working Politically' and a realisation that linear models of change apply poorly to politically influenced governance and human rights work. This re-design has involved developing a new approach to M&E, whereby Outcome Harvesting has become our main annual data collection methodology. Outcomes harvested are used for ongoing reflection about whether the programme strategies as set out in the Theory of Change are working. This process, called 'Strategy Testing', allows us to respond flexibly to changes in the context, but more importantly, it allows for critical reflection when a strategy does not seem to be producing the results expected and should therefore be stopped, or adapted to be more likely to be effective. The session will explore this process and the role Outcome Harvesting plays in it.

Thursday, 4 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 32 Using Innovative Geospatial Analysis to Evaluate Geographic Footprint of Development Interventions****Using Innovative Geospatial Analysis to Evaluate Geographic Footprint of Development Interventions***Z. Bogetic<sup>1</sup>, J. Vaessen<sup>2</sup>, A. Kumar<sup>1</sup>*<sup>1</sup> IEG- The World Bank, Human development and Economic Management, Washington D.C., USA<sup>2</sup> IEG- The World Bank, Methods team, Washington D.C., USA

**Rationale:** The World Bank established overarching twin goals in 2013: ending poverty by 2030 and sharing prosperity. The first goal is focused on the reduction of global extreme poverty from 10.7 percent in 2013 to 3 percent by 2030, based on the international poverty line of \$1.90 per person in 2011 purchasing power parity U.S. dollars. The second goal is a new one. Its basic metric is growth in the real incomes of the bottom 40 percent of the income distribution of the population (the bottom 40) in each country. Because of limitations in spatial distributional data, however, little is known about geographical allocation among the institution's projects within countries as well as their congruence with the distribution of the bottom 40 percent, for example. The same holds for other bilateral and multilateral development actors. Nonetheless, because of the recent increase in the availability of disaggregated spatial data, measuring and analyzing the subnational distribution of project funding and its relationship to the bottom 40 are now easier. Better understanding of the geographic footprint can improve efficiency and help maximize the poverty-reducing and broader development effects of development programs.

**Objective:** Present and discuss with evaluator audience innovative evaluative research using geospatial analysis of World Bank projects and their congruence with the bottom 40 percent of the populations in 58 countries as well as a geospatial country case study of Mexico.

**Panel session narrative:** This session presents and discusses two innovative evaluative papers that provide first IEG geospatial analyses of the geographic patterns of World Bank projects at the subnational level, their correlation with the spatial distribution of the bottom 40 percent, as well as the factors behind the observed correlations. The first paper "Putting your money where your mouth is: Geographic targeting of World Bank projects to the bottom 40 percent," presented by Bogetic, provides a comprehensive analysis of the congruence between the spatial distribution of the World Bank investment projects at subnational level and the bottom 40 percent in 58 countries and employs regression analysis to understand the factors behind the observed correlations using fixed effects and controls for relevant variables as well as a number of robustness tests. The paper finds that there is no clear evidence that the projects spatially target the bottom 40 percent and that there is a "capital city effect" whereby projects often locate in a region with the capital city. The paper provides recommendations for future Bank monitoring of the location of projects in relation to the bottom 40 percent. The second paper, "Geospatial analysis of World Bank projects in Mexico," presented by Kumar, applies the geospatial methodology in analyzing the spatial distribution of World Bank projects in a large upper middle-income country, Mexico, and its congruence with the bottom 40 percent. The paper extends the analysis to include knowledge work of the Bank, which often accompanies projects as part of clusters of development interventions. The panel will also discuss issues in geospatial data, geocoding and in establishing causality.

**PN 32 Using Innovative Geospatial Analysis to Evaluate Geographic Footprint of Development Interventions****PC 124 - Putting Your Money where Your Mouth is: Geographic Targeting of World Bank Projects to the Bottom 40 Percent**

*Z. Bogetic<sup>1</sup>, H. Öhler<sup>2</sup>, M. Negre<sup>3</sup>, R. Massari<sup>4</sup>, L. Smets<sup>5</sup>*

<sup>1</sup> IEG- The World Bank, Human development and Economic Management, Washington D.C., USA

<sup>2</sup> German Development Institute, Researcher, Bonn, Germany

<sup>3</sup> German Development Institute/World Bank, Senior Researcher, Bonn, Germany

<sup>4</sup> World Bank, Consultant, Washington, DC, USA

<sup>5</sup> Inter-American Development Bank, Economics Senior Specialist, Port of Spain, Trinidad and Tobago

It provides a comprehensive analysis of the congruence between the spatial distribution of the World Bank investment projects at subnational level and the bottom 40 percent in 58 countries and employs regression analysis to understand the factors behind the observed correlations using fixed effects and controls for relevant variables as well as a number of robustness tests. The paper finds, inter alia, that after appropriate controls there is no clear evidence that the Bank's investment projects spatially target the bottom 40 percent. There is also evidence of the capital city effect where projects often locate in a region with the capital city. The paper recommends that the World Bank systematically collect relevant data and monitor the spatial congruence between its projects and the bottom 40 percent in the future.

**PN 32 Using Innovative Geospatial Analysis to Evaluate Geographic Footprint of Development Interventions****PC 125 - Geospatial Analysis of World Bank Projects in Mexico**

*A. Kumar<sup>1</sup>, Z. Bogetic<sup>1</sup>, M. Negre<sup>2</sup>, H. Öhler<sup>3</sup>*

<sup>1</sup> IEG- The World Bank, Human Development and Economic Management, Washington D.C., USA

<sup>2</sup> German Development Institute/World Bank, Senior Researcher, Bonn, Germany

<sup>3</sup> German Development Institute, Researcher, Bonn, Germany

by Negre, Öhler, Bogetic and Kumar, is presented by Kumar. The paper applies the geospatial methodology in analyzing the spatial distribution of World Bank projects in a large upper middle-income country, Mexico. All World Bank projects attributable to Mexican states in the period 2008–2017 are subject of the geospatial analysis. The paper extends the analysis to include knowledge work of the Bank, which often accompanies projects. The paper provides answers to the following questions: what is the correlation between the geographic footprint of Bank projects (and knowledge products) in Mexico and the geographical distribution of the bottom 40? Does the government's own public funding correlate with the bottom 40 percent? And what factors may influence the spatial allocation of World Bank interventions in Mexico?

Thursday, 4 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 33 Assessment Methodologies for Transformational Governance and Community Resilience: Applications in Gang Violence Prevention, Countering Violent Extremism and Resilient Governance Systems**

## **Assessment Methodologies for Transformational Governance and Community Resilience: Applications in Gang Violence Prevention, Countering Violent Extremism and Resilient Governance Systems**

*M.B. Palmisano*<sup>1</sup>

<sup>1</sup> *Creative Associates International, Design Monitoring and Evaluation, Washington DC, USA*

Creative Associates International (Creative) has been implementing programs aimed at strengthening community resilience in conflict and post-conflict contexts for 40 years. Central to Creative's approach to building resilient societies is the foundational notion that resilience is not merely the ability to bounce back from shocks, but rather a set of qualities that transform rules, roles, and responsibilities within a societal system so that conditions that led to fragility in the first place are not recreated. To this end, Creative has developed and adapted assessment methodologies that can capture individual, institutional, and systemic transformations and has built local capacities to utilize these methodologies to adapt context-appropriate interventions while leveraging local strengths, and track progress towards resilience. This panel will discuss the application of these methodologies to measure resilience outcomes for three thematic areas: gang violence prevention, countering violent extremism (CVE), and promoting resilient governance systems. The session will first cover the Youth Services Eligibility Tool (YSET). YSET is a diagnostic tool to assess youth risk levels and vulnerability to crime and violence in areas where violence and conflict are chronically prevalent. The tool was adopted from the Gang Reduction and Youth Development program in Los Angeles, USA. Creative has adapted and applied YSET in Honduras, El Salvador, and the Caribbean to prevent gang violence by reducing risk factors and reinforcing family protective factors. Through YSET, youth are evaluated against nine risk factors. In Tunisia, YSET has been adapted to address the behavior of youth at high risk of joining violent extremist organizations and was complemented by in-depth qualitative studies to determine risk factors and identify required interventions. Second, Creative's Governance Fragility Resilience Assessment Method (FRAME) will be presented. FRAME is a citizen-inclusive participatory approach to assessing the resilience of governance systems in complex environments. FRAME uses a "whole-of-society" approach to designing and assessing transformational governance by recognizing dynamic interconnections between governance actors, including local councils, private sector, civil society, and citizens affiliating with different identity groups. Given that these actors may value and experience these interconnections differently, FRAME engages them in "rating" their community along eight functional dimensions. This exercise amplifies traditionally excluded voices and places them at the center of the evaluation process. Finally, the session will cover a three-phased diagnostic assessment that Creative designed and implemented for a CVE program targeting five regions in Tanzania. The tool adapts the Inter-agency Conflict Analysis Framework (ICAF) – a US government interagency conflict assessment tool – to examine drivers of violent extremism (VE) and inform the development of preventative responses that capitalize on regional resiliencies. The tool bolsters the understanding of interconnected VE dynamics and identifies potential feedback loops that may influence VE drivers through engaging research and evaluation teams in a participatory systems-mapping process looking at community relationships and behaviors. Following the mapping exercise, Creative works with communities to develop community-based theories of change to prevent the rise or spread of VE.

**PN 33 Assessment Methodologies for Transformational Governance and Community Resilience:  
Applications in Gang Violence Prevention, Countering Violent Extremism and Resilient Governance  
Systems**

**PC 126 - Measuring Individual and Community Resilience:  
Applications to Gang Violence in Central America and CVE  
in Tunisia**

*M.B. Palmisano*<sup>1</sup>

<sup>1</sup> *Creative Associates International, Design Monitoring and Evaluation, Washington DC, USA*

Building the resilience of individuals, families and communities is one of the major stated objectives of development interventions especially in areas where violence and conflicts are chronically prevalent. Creative Associates has been implementing programs with the objective of building resilience at the individual, family, and community levels for the past few decades. The Youth Services Eligibility Tool (YSET) is a diagnostic, monitoring and evaluation tool to assess youth risk levels and vulnerability to crime and violence. The tool was adopted from the Gang Reduction and Youth Development program in Los Angeles, USA in the early 2000s. Creative has adapted and applied YSET in Honduras, El Salvador, and the Caribbean to prevent gang violence by reducing risk factors and reinforcing family protective factors. Through YSET, youth are evaluated against nine risk factors/domains. At baseline, youth risk levels are categorized into low, moderate and high. In Honduras, El Salvador, and the Caribbean, working with families with 8 to 17-year-olds, in dangerous municipalities, the interventions aim to reduce risk factors to violence and criminal behavior. Using a risk-differentiated approach, youth are then referred to interventions that address their individual risk levels. Follow-up assessments are implemented at six months intervals to measure impact and track progress over time. In Tunisia, YSET has been adapted to address the behavior of youth at high risk of joining violent extremist organizations and was complemented by in-depth qualitative studies on the local context to determine risk factors and identify required interventions. Follow-up assessments after 12 months has been conducted to measure change from the baseline to guide further actions. Early indications are that while the tool is promising to measure resilience there are important challenges and lessons to consider in collection, analysis and interpretation of perception-based composite indicators from survey data. For best results, combining perception-based quantitative survey with qualitative component allow to contextualize the quantitative findings.

**PN 33 Assessment Methodologies for Transformational Governance and Community Resilience:  
Applications in Gang Violence Prevention, Countering Violent Extremism and Resilient Governance  
Systems**

**PC 127 - Fragility Resilience Assessment Method (FRAMe):  
An Approach to Designing and Assessing Transformational  
Governance Resilience Programs**

*M. Proctor*<sup>1</sup>

<sup>1</sup> *Creative Associates International, Technical Manager with the Governance and Community Resilience practice area, Washington DC, USA*

Creative Associates' Governance Fragility Resilience Assessment Method (FRAMe) is a citizen-inclusive participatory approach to assessing the resilience of governance systems in complex environments. FRAMe was initially piloted through Creative's pioneering work with newly emergent citizen councils in Syria. FRAMe uses a "whole-of-society" approach to designing and assessing transformational governance programs by recognizing the dynamic interconnections between different governance actors, including local councils, private sector, civil society, and citizens affiliating with different identity groups. Given that these actors may perceive, experi-

ence, and value these interconnections differently, FRAMe assesses the resilience of a governance system by engaging cross-sections of these governance actors directly in a participatory self-assessment. FRAMe accounts for governance actors' different perspectives on the relevant roles, relationships, and responsibilities within the governance system. Using FRAMe, governance actors rate eight dimensions of their governance system. Each dimension is assessed according to seven factors:

- (1) inclusion;
- (2) decentralization/local discretion;
- (3) social cohesion;
- (4) performance;
- (5) civic infrastructure;
- (6) confidence and trust; and
- (7) system legitimacy.

This exercise elevates and amplifies voices that are traditionally excluded from community-level problem-solving – such as those of young people, women, and ethnic minorities – and places them at the center of the evaluation process. The exercise also serves to elucidate key grievances and conflict drivers, highlight cleavages between identity groups, and map possible trajectories towards more resilient outcomes. Through FRAMe, target communities are better equipped to leverage local strengths, adapt and adjust programs and resources, and develop paths to resilience while tracking their progress on the fragility-resilience continuum. The panel discussion will further explain the FRAMe approach and present results from its application in the newly emergent citizen councils in Syria (2016–2017), its field-testing in Mali (2017), and the most recent FRAMe implementation in El Salvador (2018).

Thursday, 4 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 34 The Economics of Resilience Returns: When 1 + 1 is Greater Than 2**

## The Economics of Resilience Returns: When 1 + 1 is Greater Than 2

B. Sagara<sup>1</sup>

<sup>1</sup> Mercy Corps, Research and Learning, Oakland, USA

It is believed that investing in resilience, preparedness, and early response can lead to savings in lives and livelihoods preserved, humanitarian cost avoided, and development gains sustained. However, evidence testing this assumption is limited. Filling this gap is critical as cost analyses can provide information to decision makers on how to maximize every dollar spent to improve lives. The objectives of this panel are to: 1) Present examples of cost analyses of resilience building interventions, 2) Discuss promising practices and challenges in these analyses to improve evaluation practice, 3) Appreciate the donor perspective in the utilization of these evaluations to inform policy and investment strategies; identify evidence gaps that evaluators should be focusing on. This panel presents four unique perspectives on conducting and using cost analysis in evaluation of resilience investments. Courtenay Cabot-Venton and Gil Yaron have conducted cost analyses in East Africa and Myanmar respectively and offer distinct perspectives in terms of methods and scale; with Courtenay focusing at a regional level and Gil focusing at the sub-national project level. Nyoman Prayoga from Mercy Corps Indonesia will share lessons learned from the practitioner perspective on the use of these analyses. Tiffany Griffin from the USAID Center for Resilience provides insight into how these analyses inform and guide decision-making from the donor perspective.

**PN 34 The Economics of Resilience Returns: When 1 + 1 is Greater Than 2**

## PC 128 - Evaluating the Returns to Community-Based Climate Resilience Interventions

G. Yaron<sup>1</sup>

<sup>1</sup> Gil Yaron Associates, Harpenden, United Kingdom

Recent work by Gil supported by the DFID BRACED programme and the Asian Development Bank has provided evidence for community resilience interventions that combine rigorous participatory methods with robust economic estimates of impacts such as loss of life in Myanmar. This study was used to calculate 'resilience dividends' – those beneficial impacts of resilience programming such as avoided losses and spin off health benefits such as reduction in dengue incidence.

**PN 34 The Economics of Resilience Returns: When 1 + 1 is Greater Than 2**

## PC 129 - Rising Every Time We Fall: The Economics of Resilience

L. Murphy Michalopoulos<sup>1</sup>

<sup>1</sup> USAID, Center for Resilience, Washington- DC, USA

Tiffany will discuss the evolution and importance of the economic rationale for resilience from the donor perspective. Two approaches for identifying the value of resilience strengthening will be discussed – a scenario-based modeling approach and an approach using actual empirical data. Policy implications in a climate of ever-shrinking resources, as well as programmatic ramifications for resilience projects will be examined.

Thursday, 4 October 2018  
08:30–10:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 35 Using Monitoring and Evaluation to Help Societies Become More Resilient**

## Using Monitoring and Evaluation to Help Societies Become More Resilient

*P. Van Nierop*<sup>1</sup>

<sup>1</sup> ICF, European Policy, London, United Kingdom

In the last years, several conflicts, disasters and/or humanitarian emergencies have happened that needed an immediate, coordinated and effective response to protect the most vulnerable and avoid wider effects in Europe and abroad. To mention a few: the Syrian conflict affected over 12 million people within the country, half of which were children; the Sahel hunger crisis left over 20 million suffering food insecurity and the Ukraine conflict led to the displacement of over 2 million people. These crises are also linked to large migratory movements to Europe of people fleeing conflicts and disasters which need an immediate and sustainable response. Against this background, the question of how to make societies in Europe and abroad more resilient is becoming more pressing than ever. The panel will explore how monitoring and evaluation can promote resilience and action in turbulent times. It will discuss how monitoring and evaluation can increase the capacity of societies to respond to a humanitarian crisis or emergency, by being better prepared, improving the quality of decision-making, ensuring the relevance of actions, dealing more effectively with the aftermath of a crisis and ensuring sustainability. It will also discuss the challenges of monitoring and evaluation in disasters and humanitarian emergencies. The panel will present different types of monitoring and evaluation approaches, which include: (i) Setting up monitoring systems to better understand the scale and scope of a crisis, to develop the most appropriate response and to be able to measure its effects; (ii) Evaluating the processes put in place to respond to humanitarian emergencies to ensure that they can be adjusted 'real time'; (iii) Evaluating ex post the treatment of victims of a crisis (such as IDPs, migrants and refugees) to ensure that any coordinated actions introduced at European or international level were effective, efficient and provided added value; (iv) Evaluating resilience. Panelists will present concrete examples of responses to disasters or humanitarian emergencies and assess opportunities and challenges of monitoring and evaluation in each case. Regarding monitoring, examples can include the monitoring of reception capacity for asylum seekers and the monitoring of integration of refugees. In relation to evaluations, examples could include evaluating the application of EU law on criteria to granting international protection and rights to asylum seekers and evaluating whether actions for victims are needs based and effective, as well as support resilience (all humanitarian aid evaluations). Some of the key questions that the panel will tackle through the presentations and discussions are: (i) Are there any specific evaluation designs, approaches and methodologies that work particularly well? (ii) What are the (proven) benefits of the types of approaches that are being discussed? (iii) What are the challenges and constraints? (iv) What further improvements could be introduced to enhance the usefulness of monitoring and evaluation?

**PN 35 Using Monitoring and Evaluation to Help Societies Become More Resilient**

## PC 130 - The Role of Monitoring and Evaluation in DG ECHO's Interventions Relating to Resilience Building and Disaster Response

*J. Nilsson*<sup>1</sup>

<sup>1</sup> European Commission – Directorate-General for European Civil Protection and Humanitarian Aid Operations (DG ECHO), Brussels, Belgium

This presentation discusses:

- 1) DG ECHO's monitoring tools for planning, defining, and implementing actions;
- 2) DG ECHO's evaluation framework and programme, including practical examples;

and how the above activities contribute to ensuring quality of actions and building resilience of societies.

#### **PN 35 Using Monitoring and Evaluation to Help Societies Become More Resilient**

### **PC 131 - Challenges to Evidence Quality in Evaluations of Humanitarian Action**

*N. Dillon*<sup>1</sup>

<sup>1</sup> ALNAP, London, United Kingdom

This presentation discusses the strengths and weaknesses of evaluations of humanitarian action, through the prism of a sector-wide evaluation synthesis that fed into the 2018 edition of the ALNAP State of the Humanitarian System Report. The presentation provides an assessment of the evidence quality generated by humanitarian evaluations across the critical thematic areas of relevance/appropriateness, coverage and effectiveness, as well as exploring the types of evaluation methods most commonly used and their relationship to evidential quality. Driving factors for challenges to evidence quality are considered, including humanitarian sector-specific implementation constraints.

#### **PN 35 Using Monitoring and Evaluation to Help Societies Become More Resilient**

### **PC 132 - Can Monitoring Help European Societies to Become More Resilient to Migratory Movements and Protect the Most Vulnerable?**

*K. Mantouvalou*<sup>1</sup>

<sup>1</sup> ICF, London, United Kingdom

This presentation will explore how monitoring can help European Member States respond more effectively to large migratory movements and protect the most vulnerable. Taking as case studies three recent examples of monitoring of

- a) the capacity of reception centres,
- b) asylum procedures and
- c) the health of newly arrived asylum seekers,

the presentation will discuss how monitoring can support societies to improve the quality of decision-making, ensure the relevance of actions, deal more effectively with the aftermath of a crisis and ensure sustainability. It will also discuss the challenges of monitoring in humanitarian emergencies.

Thursday, 4 October 2018  
11:30–13:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### PN 36 Addressing Ethics and Values in Evaluation

## Addressing Ethics and Values in Evaluation

*N. Stame*<sup>1</sup>

<sup>1</sup> *University of Rome "La Sapienza", communication and social research, Roma, Italy*

The Conference title evokes turbulent times characterized by negative traits: inequality, populism, nationalism, environmental damage, etc. It invites inquiry into the role that evaluation could play in supporting policies that could counteract those tendencies. To this end, it anticipates the capacity of societies under stress to be resilient – to create new opportunities for social wellbeing and development from their own potential.

This would imply that evaluation overcome some of its main difficulties in dealing with ethical issues.

Up to now most efforts have been directed towards the ethical behavior of evaluators often ignoring the value implications and even intention of what is being evaluated. In this panel we are interested in how evaluation and evaluators might move beyond the instrumental identification with pre-set goals, with success/failure, and take account of the values inherent in programs and policies and the beliefs and motivations of their actors. We hope, in part and ambitiously, to explore whether and how it might be possible to use as evaluative criteria 'what is good and what is bad for societies and for people'.

The panel will explore issues involved in doing value oriented evaluations and argue the need for evaluation that is value oriented. Its objective is to identify lessons that evaluation could learn from other knowledge sources; Elaborate criteria and research methods that could suit value oriented evaluations.

### PN 36 Addressing Ethics and Values in Evaluation

## PC 133 - Addressing the Difficulty for Value-Oriented Evaluations: Could a Moral Social Science Offer an Ethically Relevant Perspective?

*N. Stame*<sup>1</sup>

<sup>1</sup> *University of Rome "La Sapienza", communication and social research, Roma, Italy*

Evaluators are familiar with ethical issues related to "acting ethically", much less with "analyzing ethically", thanks to the still prevalent mantra of the separation between facts and values that is supposed to regulate evaluation research. This presentation looks for streams of social science that consider morality as a crucial dimension of analysis, and that aim at providing empirical evidence morally significant. Such approaches could offer a favorable self-reassurance for evaluators critical of the mainstream, who are concerned with the current challenges of growing inequality, social crises and institutional weakness that social policies find difficult to counteract.

**PN 36 Addressing Ethics and Values in Evaluation****PC 134 - Valuing Without Shared Values**

*E. Stern*<sup>1</sup>

<sup>1</sup> *University of Lancaster, emeritus professor, Lancaster, United Kingdom*

Constructionist and deliberative scholars suggest that evaluators having identified values in unbiased ways, should aim to empower the excluded; and/or promote consensus amongst those with diverse values. Participation is itself a 'good' in a pluralist society. This stance assumes 1) a largely 'vertical' or power or inequality-based classification of value differences; and 2) higher level values that are inclusive and to which all can subscribe. What's to be done when value differences are horizontal as well as vertical; and when we can't always appeal to shared values?

**PN 36 Addressing Ethics and Values in Evaluation****PC 135 - Ethical Justifications for Collaborative Approaches to Evaluation**

*B. Cousins*<sup>1</sup>

<sup>1</sup> *University of Ottawa, emeritus professor, Ottawa, Canada*

Collaborative approaches to evaluation (CAE) necessarily imply partnership between evaluators and members of the programme community (stakeholders, broadly defined) to jointly produce evaluative knowledge. To date, CAE has relied on three principal justifications: PRAGMATIC or practical, problem-solving motives; POLITICAL or transformative/emancipatory objectives; and PHILOSOPHICAL or epistemological reasons relating to the development of deeper understandings of complex phenomena. In the face of growing moral-political challenges facing societies a fourth rationalisation for CAE requires serious consideration. This presentation will explore ETHICAL justifications for CAE from the perspective of values and beliefs inherent in programmes and programme communities. Some key questions are: Who among the stakeholders is doing the valuing and how are evaluators implicated? What (if any) is the moral impetus for the evaluation? What is the moral reasoning shaping methodological and relational choices within the evaluation?

**PN 36 Addressing Ethics and Values in Evaluation****PC 136 - Valuing in the Service of the public Interest(s) in Our Complex World**

*G. Julnes*<sup>1</sup>

<sup>1</sup> *University of Baltimore, School of Public and International Affairs, Baltimore, USA*

To the extent that evaluation, as a field and a practice, has a moral imperative, it is often expressed as "contributing to the improvement of the program or policy" (Weiss, 1998, p. 4), which for social programs involves the "aim to improve the welfare of individuals, organizations, and society" (Shadish, Cook, & Leviton, 1991, p. 91). That there is no consensus on what the "improvement of welfare" means, much less on how to assess it, is increasingly problematic given (1) the increasing expectation of "evidence-based policymaking," meaning that misunderstanding "societal welfare" has greater potential for harm, and (2) the emerging understanding of the complexity of the processes that programs and policies are expected to manage. This presentation will address the challenges and opportunities for better assessment of the "value" of programs and policies and also what this means in the context of sustainability.

**PN 36 Addressing Ethics and Values in Evaluation****PC 137 - The West and the Rest: an Exploration of the Role of Societal Values in Shaping Evaluation for Development**

Z. Ofir<sup>1</sup>

<sup>1</sup> University of Stellenbosch, School of Public Leadership, Cape Town, South Africa

Both 'development' and of 'evaluation' have been shaped over the last six decades by a limited number of dominant paradigms. These ways of viewing the world have their origin primarily – although not exclusively – in economics (in the case of development) and various other social sciences (in the case of evaluation) in line with how these fields have evolved in the West. Although the 2030 Agenda for Sustainable Development and impressive development successes outside the norm have brought new perspectives to the fore, evaluators still struggle to move beyond the conventional mental models that frame or influence our theories and practices. What is the role of societal values in shaping this phenomenon?

**PN 36 Addressing Ethics and Values in Evaluation****PC 138 - The Normative Political Characteristics of Professional Ethics in Evaluation: the Case for Democratic Professionalism**

T. Schwandt<sup>1</sup>

<sup>1</sup> University of Illinois, emeritus professor, Urbana Champaign III, USA

What ethics means in the field of evaluation is largely confined to matters of face-to-face interaction of professionals with those with whom professionals work; that is, what is commonly referred to as professional ethics. Less attention is given to the normative characteristics that are unique to evaluation professionalism. This paper focuses on the normative political characteristics of professional ethics in evaluation, that is, how the profession ought to be connected to conceptions of the citizenry and the common good. It argues for a professional ethic referred to as democratic professionalism.

Thursday, 4 October 2018  
11:30–13:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 37 Learning through Principles Discovery Methods When the Program Theory is Unclear: A Data-Driven Approach to Understanding Context-Mechanisms-Outcomes Associated with Impacts**

## Learning through Principles Discovery Methods when the Program Theory is Unclear: A Data-Driven Approach to Understanding Context-Mechanisms-Outcomes Associated with Impacts

*S. Sridharan*<sup>1</sup>, *M. Mark*<sup>2</sup>, *A. Dey*<sup>3</sup>, *A. Nakaima*<sup>4</sup>, *S. Krishnan*<sup>5</sup>

<sup>1</sup> *University of Toronto and St. Michaels Hospital, The Evaluation Centre for Complex Health Interventions, Toronto- Ontario, Canada*

<sup>2</sup> *Penn State University, Psychology, University Park, USA*

<sup>3</sup> *Sambodhi Research and Communication Limited, Sambodhi, Noida-, India*

<sup>4</sup> *St. Michaels Hospital, The Evaluation Centre for Complex Health Interventions, Toronto, Canada*

<sup>5</sup> *Gates Foundation, Monitoring- Learning and Evaluation, New Delhi, India*

This panel will explore the role of methods when faced with theories of change that are incomplete. We will explore how the methods themselves can be useful to build and refine theories of change over time as well as learn about programme impacts over time. Of specific interest in this panel will be a range of complex interventions all with varying degrees of complexity. The interventions vary from a dance intervention focused on Parkinson's Disease in Toronto, Canada; a maternal health intervention in the Indian state of Uttar Pradesh; and systems of interventions that comprise the Sustainable Development Goals. Different paradigms of learning about refining program theories and impacts over time will be explored. In addition, the notion of incompleteness in our understanding of interventions and their theories of change along with incompleteness in knowledge of what contexts might matter will be explored. The relevance of methods to identify contexts that matter, as well as how these methods themselves can precipitate thinking about differential programme mechanisms in different contexts will be explored.

The panel will consist of the following speakers: Arnab Dey, Suneeta Krishnan, April Nakaima, Sanjeev Sridharan and Mel Mark. Arnab and Suneeta will discuss how relevant contexts that might matter were identified in a maternal health intervention in the Indian state of Uttar Pradesh. April will discuss the utility of Principled Discovery methods in the Dancing With Parkinson's intervention. Sanjeev will connect these ideas with notions of intersectionality when the theories of change are limited. Mel will discuss the relevance of Principled Discovery methods when the knowledge of programme theories is incomplete. Both methods of Principled Discovery and Competitive Elaboration will form the focus of the discussion.

The implications of such a panel include: how implementation of programmes need to be multi-phased with an initial phase being learning about contexts and intersections of context that might matter; a second implication will be the role of methods to learn about programme mechanisms as programmes get implemented; third, the utility of methodological paradigms for learning about Principled Discovery methods in learning about heterogeneous impacts will be explored.

**PN 37 Learning through Principles Discovery Methods When the Program Theory is Unclear:  
A Data-Driven Approach to Understanding Context-Mechanisms-Outcomes Associated with Impacts**

**PC 139 - Applications of Principled Discovery Methods when  
the Theory of Change of Interventions is Unclear**

S. Sridharan<sup>1</sup>, M. Mark<sup>2</sup>, A. Dey<sup>3</sup>, A. Nakaima<sup>4</sup>, S. Krishnan<sup>5</sup>

<sup>1</sup> University of Toronto and St. Michaels Hospital, The Evaluation Centre for Complex Health Interventions, Toronto- Ontario, Canada

<sup>2</sup> Penn State University, Psychology, University Park, USA

<sup>3</sup> Sambodhi Research and Communication Limited, Sambodhi, Noida-, India

<sup>4</sup> St. Michaels Hospital, The Evaluation Centre for Complex Health Interventions, Toronto, Canada

<sup>5</sup> Gated Foundation, Monitoring- Learning and Evaluation, New Delhi, India

The ideas discussed in this presentation find resonance in the ideas of what Mark, Henry and Julnes (2000) call principled discovery. Mark et al (1998, 14) pose the question, "How do we ask the data, rather than practitioners or social science theory, to provide the program theory to further guide us?" Mark, Henry and Julnes (2000, p. 259) described principled discovery as a method that, "...can allow discovery via induction within the complexities of an open system but that are principled in that the discoveries are subsequently disciplined by data and are not simply post hoc explanations that exploit chance variations in a particular sample." Basic ideas of principled discovery methods will be presented.

**PN 37 Learning through Principles Discovery Methods When the Program Theory is Unclear:  
A Data-Driven Approach to Understanding Context-Mechanisms-Outcomes Associated with Impacts**

**PC 140 - Understanding and Identifying Relevant Contexts  
in Maternal Health Settings in India**

S. Sridharan<sup>1</sup>, M. Mark<sup>2</sup>, A. Dey<sup>3</sup>, A. Nakaima<sup>4</sup>, S. Krishnan<sup>5</sup>

<sup>1</sup> University of Toronto and St. Michaels Hospital, The Evaluation Centre for Complex Health Interventions, Toronto- Ontario, Canada

<sup>2</sup> Penn State University, Psychology, University Park, USA

<sup>3</sup> Sambodhi Research and Communication Limited, Sambodhi, Noida-, India

<sup>4</sup> St. Michaels Hospital, The Evaluation Centre for Complex Health Interventions, Toronto, Canada

<sup>5</sup> Gates Foundation, Monitoring- Learning and Evaluation, New Delhi, India

This presentation will discuss how methods can be used to identify contexts that matter. Using a mixed methods design, we will discuss how relevant contexts can be identified using an exploratory approach and how heterogeneous impacts can be estimated using a combination of principled discovery and design-focused methods. This presentation will draw on a complex intervention that has multiple components. The utility of principled discovery approaches when faced with complex interventions will be highlighted.

**PN 37 Learning through Principles Discovery Methods When the Program Theory is Unclear:  
A Data-Driven Approach to Understanding Context-Mechanisms-Outcomes Associated with Impacts**

**PC 141 - Applying Principled Discovery Approaches  
to Evaluating the Impacts of a Dance-based Intervention  
Focused on Clients with Parkinson's Disease**

S. Sridharan<sup>1</sup>, M. Mark<sup>2</sup>, A. Dey<sup>3</sup>, A. Nakaima<sup>4</sup>, S. Krishnan<sup>5</sup>

<sup>1</sup> University of Toronto and St. Michaels Hospital, The Evaluation Centre for Complex Health Interventions, Toronto- Ontario, Canada

<sup>2</sup> Penn State University, Psychology, University Park, USA

<sup>3</sup> Sambodhi Research and Communication Limited, Sambodhi, Noida-, India

<sup>4</sup> St. Michaels Hospital, The Evaluation Centre for Complex Health Interventions, Toronto, Canada

<sup>5</sup> Gates Foundation, Monitoring- Learning and Evaluation, New Delhi, India

This presentation will describe a mixed methods design that was informed by a realist approach to understand the impacts of a dance intervention in Toronto, Canada. Using a longitudinal design, we demonstrate how a theory-driven evaluation approach that informed a principled discovery framework can be integrated within a longitudinal design to understand how a dance intervention works for participants affected to varying degrees by the movement and balance challenges associated with Parkinson's Disease.

**PN 37 Learning through Principles Discovery Methods When the Program Theory is Unclear:  
A Data-Driven Approach to Understanding Context-Mechanisms-Outcomes Associated with Impacts**

**PC 142 - Taking Intersectionalities Seriously: The Role  
of the Importance of Identifying Multiple Intersecting Contexts  
in Addressing Inequities**

S. Sridharan<sup>1</sup>, M. Mark<sup>2</sup>, A. Dey<sup>3</sup>, A. Nakaima<sup>4</sup>, S. Krishnan<sup>5</sup>

<sup>1</sup> University of Toronto and St. Michaels Hospital, The Evaluation Centre for Complex Health Interventions, Toronto- Ontario, Canada

<sup>2</sup> Penn State University, Psychology, University Park, USA

<sup>3</sup> Sambodhi Research and Communication Limited, Sambodhi, Noida-, India

<sup>4</sup> St. Michaels Hospital, The Evaluation Centre for Complex Health Interventions, Toronto, Canada

<sup>5</sup> Gates Foundation, Monitoring- Learning and Evaluation, New Delhi, India

This paper will discuss exploratory multilevel approaches that can be used to identify individuals who live at the intersections of multiple margins of disadvantage. One of the features of this presentation is that multilevel factors are considered in the identification of such intersections. Implications for program planning and developing theories of change are considered

Thursday, 4 October 2018  
11:30–13:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 38 Privacy by Design: Principles to Balance Impact Measurement with Privacy, Security and Safeguarding**

## Privacy by Design: Principles to Balance Impact Measurement with Privacy, Security and Safeguarding

*M. Khan*<sup>1</sup>

<sup>1</sup> *Daira, Principal, London, United Kingdom*

This panel will explore the thorny oxymorons that our sector routinely wrestles with as it integrates digital measurement with evaluation practices. How can we both enable greater voice and participation while still ensuring that we are protecting children and their privacy? How can we collect the data needed to be able to adapt platforms and content to users' needs and interests while minimizing the data we collect so that we're respecting privacy and reducing potential for data breaches? How can we use digital data to understand the contribution of mobile and online platforms to behavior change whilst complying with privacy laws and regulations? Using a 'privacy by design' approach that builds in behavior change communication goals and ways to measure them whilst also embedding safeguarding, privacy, and security mechanisms can enable us to address the complex challenges of working on digital platforms. As the sector increasingly works in the digital space, it needs to resolve the tensions around privacy, work that aims to change social norms through mobile content, and measurement. It is critical to collect data in order to adapt and enhance the content we are producing based on feedback from users of a platform and to measure behavior changes and impact, yet we also need to ensure that we are responsible about how we use social media and how we collect and use data so that we do not put vulnerable people and groups, in our case, adolescent girls, at risk. This panel will draw on digital safeguarding and digital measurement expertise to illustrate the considerations that are made 'in real time' to develop and a successful and safe digital behaviour change communications initiative. The panelists will cover the design, content and moderation cycles that contribute to a live platform, while focusing on the elements of safeguarding and privacy/security with monitoring, evaluation, research and learning. We discuss how these different and distinct elements need to work in delicate balance to ensure that a platform is safe, effective and adaptive. We outline the challenges in attaining this balance and offer tips and suggestions for how to do so. Objectives of the panel include: 1) Increased understanding of the implications of data privacy and security in digital data environments and 2) Improved awareness of methods used to responsibly collect digital data

**PN 38 Privacy by Design: Principles to Balance Impact Measurement with Privacy, Security and Safeguarding**

## PC 143 - Live Data in Impact Assessment: Privacy and Security Considerations

*K. Bertermann*<sup>1</sup>

<sup>1</sup> *Girl Effect, Evidence, London, United Kingdom*

Web-based entertainment education platforms offer unique opportunities in terms of data gathering and analysis. The content in such a platform is static and is created according to robust research related to challenges and opportunities for adolescent girls. At the same time, the platform itself is dynamic, as users interact with the content, post comments, share ideas and form mini communities in real time. A measurement framework for a digital platform, therefore, must be designed in order to gather data against static behaviour change outcomes, while simultaneously utilising live feedback and engagement data to make sense of content resonance and proxies for long term behaviour change. This combination of static and dynamic outcomes and data requires special considerations for impact assessment. As we developed

our measurement frameworks for Girl Effect's digital platforms, we drew from emerging methods in data science that are typically applied to commercial platforms, as well as traditional evaluation methods, realising that the amount and type of data from the platform enables us to understand change in new ways. Impact measurement of a digital platform thus draws on new and exciting methods as well as tried-and-tested traditional techniques. In this space of digital impact measurement, Girl Effect has found it crucially important to consider privacy and safeguarding considerations in each step of impact assessment in order to provide users (in this case, adolescent girls) with a safe digital experience that provides optimal privacy whilst providing informed, active consent for data use. In this presentation, I will provide participants with the context of the digital monitoring and evaluation approach used at Girl Effect, with a focus on considerations in 'passive' and 'active' data collection. The presentation objectives will include: – Types of user data which can be collected on digital platforms – Informed consent in digital monitoring and evaluation – Ethical and safeguarding considerations Girl Effect put in place in digital monitoring and evaluation frameworks – Criteria Girl Effect uses to assess digital monitoring tools

**PN 38 Privacy by Design: Principles to Balance Impact Measurement with Privacy, Security and Safeguarding**

**PC 144 - Putting Privacy, Safeguarding and Gender at the Heart of Evaluation Design and Digital Data Processing**

L. Raftree<sup>1</sup>

<sup>1</sup> *Independent Consultant, Independent Consultant, New York City, USA*

The rapid increase in access and use of digital tools and platforms opens new opportunities for monitoring and evaluation. Digital data provides rich insights for adjusting content and approaches and to track online behaviors. However, the picture is not all rosy. Digital monitoring activities require practitioners to balance aspects that are new and unique to the digital environment. In addition to elements like inclusion and local context, evaluators need to be aware of new nuances that influence design and monitoring. We also need to take specific measures to ensure that we are not introducing risk and harm into the lives of those we are engaging, especially in the case of women and girls who experience increased levels of online and offline harassment and other kinds of risk that are often exacerbated in the digital space. Over the past 5 years, we've conducted ongoing desk research and engaged girls and boys in participatory design research to gain insight into how they think about, access and use mobile internet and online platforms. We've consulted with parents, caretakers, communities and the wider sector to draw out good practice on mobile BCC platforms and gender. Additionally, we've studied emerging legal frameworks that provide guidance on managing privacy and security. Based on the above, four years ago we developed a first draft set of guidelines for digital privacy, security and safeguarding which were integrated into our work across multiple countries and platforms. These were updated to reflect the General Data Protection Regulation (GDPR). A new child safeguarding policy was developed with a stronger integration of digital aspects. We've also developed a data policy to provide orientation on collection, storage, transmission, use, sharing, and retention/destruction of data collected via digital platforms. Overall we've adjusting practice to comply with the GDPR, in effect as of May 2018. All this has led to conversations across the organization and engagement with the wider sector around privacy by design and data security, and discussions about how GDPR, gender and digital safeguarding should be integrated into our work, including research, monitoring and evaluation. On this panel, I will share the core elements of our privacy by design approach with a focus on – Designing social media efforts that enable data collection for insights and at the same time elevate privacy and safety – Managing new data regulations in practice, including minimization of personal and sensitive data collection – Defining legal bases for data collection and Improving consent processes – Assessing partnerships with an eye to ethical data use and privacy and security – Sharing tips, tools and templates to improve data ethics, privacy, security and transparency around data practices – Experiences in rolling out data privacy, security and safeguarding policy and practice across various organizational settings

Thursday, 4 October 2018  
11:30–13:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 40 From Theory to Real World Evaluation: How Could Resilience Help Fill the Gap?**

## From Theory to Real World Evaluation: How Could Resilience Help Fill the Gap?

*C. Rodriguez-Ariza<sup>1</sup>, P. Rodriguez-Bilella<sup>2</sup>, B. Williams<sup>3</sup>, S. Vaca<sup>4</sup>, P.D.R. Stockmann<sup>5</sup>, M. Tarsilla<sup>6</sup>, M. Bustelo<sup>7</sup>, H. Stadtmueller<sup>8</sup>*

<sup>1</sup> *FAO, FAO Ethiopia Resilience team, Addis Ababa, Ethiopia*

<sup>2</sup> *University of San Juan, Evaluation, San Juan, Argentina*

<sup>3</sup> *Consultant, Systems thinking and evaluation, Auckland, New Zealand*

<sup>4</sup> *Consultant, Evaluation, Madrid, Spain*

<sup>5</sup> *Universität des Saarlandes, Centrum für Evaluation CEval, Saarbrücken, Germany*

<sup>6</sup> *Consultant, Evaluation, Paris, France*

<sup>7</sup> *Universidad Complutense de Madrid, Political Science and Public Administration, Madrid, Spain*

<sup>8</sup> *DEVAL, Evaluation Capacity Development, Bonn, Germany*

This panel is intended to foster a critical reflection among practitioners, commissioners and users on how to address the limitations that characterize evaluations in complex settings. In order to engage both seasoned and emerging evaluators in a lively interaction, presenters will encourage the audience to reflect about 5 key themes. The panel will be innovative in that it will combine the resilience theoretical frameworks currently available with the real-world experience of evaluation professionals with more than 100 years of combined evaluation experience. In doing so, the objective of the overall session will be to foster a better understanding of what it means to be resilient in evaluation (including what does or does not work for whom, under what conditions and why) and of how to make one's own practice more resilient in the future.

Despite grounding their presentations on existing theories and frameworks, panelists will provide the audience with a well-articulated and vivid description of how they have themselves developed their resilience attitudes and skills as part of their evaluation practice over time. The audience will then walk away from this session with some good tips on how to:

- (a) to complete successfully a very turbulent evaluation assignment;
- (b) to reflect upon their own practice and share the lessons learned with the rest of the community;
- (c) to identify further opportunities for building bridges between evaluative thinking and evaluative practice.

As the panel intends to be as interactive as possible, the audience will be encouraged to ask themselves the same guiding questions that the panelist will address in their own presentations. The five topics (each topic includes a few questions) will be as follows:

- (a) **GAPS in Evaluation Commissioning and their impact on the rest of the evaluation process:** How did the ToR "mess" and/or the lack of prioritization on the commissioner's part affect the project and the evaluation? For example: What are the main mistakes in terms of setting evaluation priorities?
- (b) **DISCONNECT AMONG ToR ELEMENTS:** How were the purpose, questions and the allocated time/resources disconnected from each other? Which ToR elements were the ones that complicated the evaluation endeavor the most?
- (c) **LACK OF DATA:** What specific data was impossible to collect (even if promised in the proposal)? How was the dearth of data addressed in order to complete the evaluation? gender data?
- (d) **COMMON MISTAKES IN ANALYSIS, VISUALIZATION AND COMMUNICATION:** What specific analytical techniques were missing to assess/visualize/communicate the contributions of the interventions? In what specific ways were methods not integrated at all or incorrectly integrated into the analysis, visualization or communication? gender analysis?

- (e) INABILITY TO LEARN AND USE: What are some specific examples of institutions' and organizations' inability to learn and use evaluations? Were there any surprises in the inability to learn from and use evaluations? Is there any example of repeated lesson learned or recommendations that have not contributed to any substantial change or improvement? (gender?) Lively discussion between panelists/audience will take place

**PN 40 From Theory to Real World Evaluation: How Could Resilience Help Fill the Gap?**

**PC 150 - From Theory to Real World Evaluation Capacity Development: A Transformative and Resilient Perspective from West and Central Africa**

*M. Tarsilla*<sup>1</sup>

<sup>1</sup> UNICEF- West and Central Africa Regional Office, Evaluation, Dakar, Senegal

Drawing on his extensive Evaluation Capacity Development (ECD) experience in Africa, the presenter will address a number of issues related to the issue of resilience in Evaluation Capacity Development in West and Central Africa. Consistent with the objectives of the overall panel, the presenter will, among other things, touch upon the following themes:

- (a) Main gaps in ECD Commissioning;
- (b) Disconnect between the purpose of ECD programs and the amount of resources made available;
- (c) Lack of data on the capacity of the groups targeted by ECD programs;
- (d) Weaknesses in the way evaluation topics and good practices are introduced in ECD programs. The presenter will encourage the audience to contribute their perspectives - from within and outside of Europe and Africa - on the issues touched upon during the presentation.

**PN 40 From Theory to Real World Evaluation: How Could Resilience Help Fill the Gap?**

**PC 151 - From Theory to Real World Evaluation: How Could Resilience Help Fill the Gap? Independent Evaluation, DV and Real Participation**

*S. Vaca*<sup>1</sup>

<sup>1</sup> Consultant, Evaluation, Madrid, Spain

From the perspective of the Independent evaluation, DV and real participation, Sara will address some of the following questions by providing real examples of her professional live, stories about her active evaluation experiences (empirical, self-experienced examples). Solutions will be suggested to overcome these gaps and also be thematized:

- (a) GAPS in Evaluation Commissioning and their impact on the rest of the evaluation process: How did the ToR "mess" and/or the lack of prioritization on the commissioner's part affect the project and the evaluation? For example: What are the main mistakes in terms of setting evaluation priorities?
- (b) DISCONNECT AMONG ToR ELEMENTS: How were the purpose, questions and the allocated time/resources disconnected from each other? Which ToR elements were the ones that complicated the evaluation endeavor the most?
- (c) LACK OF DATA: What specific data was impossible to collect (even if promised in the proposal)? How was the dearth of data addressed in order to complete the evaluation?
- (d) COMMON MISTAKES IN ANALYSIS, VISUALIZATION AND COMMUNICATION: What specific analytical techniques were missing to assess/visualize/communicate the contributions of the interventions? In what specific ways were methods not integrated at all or incorrectly integrated into the analysis, visualization or communication?

- (e) INABILITY TO LEARN AND USE: What are some specific examples of institutions' and organizations' inability to learn and use evaluations? Were there any surprises in the inability to learn from and use evaluations? Is there any example of repeated lesson learned or recommendations that have not contributed to any substantial change or improvement?

**PN 40 From Theory to Real World Evaluation: How Could Resilience Help Fill the Gap?**

**PC 152 - From Theory to Real World Evaluation: How Could Resilience Help Fill the Gap? Evaluation Uses / Users and Gender Gaps**

*M. Bustelo*<sup>1</sup>

<sup>1</sup> *Universidad Complutense de Madrid, Political Science and Public Administration, Madrid, Spain*

From the perspective of the evaluation use and gender mainstreaming, Maria will address some of the following questions by providing real examples of her professional life, stories about her active evaluation experiences (empirical, self-experienced examples). Solutions will be suggested to overcome these gaps and also be thematized:

- (a) GAPS in Evaluation Commissioning and their impact on the rest of the evaluation process: How did the ToR "mess" and/or the lack of prioritization on the commissioner's part affect the project and the evaluation? For example: What are the main mistakes in terms of setting evaluation priorities?
- (b) DISCONNECT AMONG ToR ELEMENTS: How were the purpose, questions and the allocated time/resources disconnected from each other? Which ToR elements were the ones that complicated the evaluation endeavor the most?
- (c) LACK OF DATA: What specific data was impossible to collect (even if promised in the proposal)? How was the dearth of data addressed in order to complete the evaluation?
- (d) COMMON MISTAKES IN ANALYSIS, VISUALIZATION AND COMMUNICATION: What specific analytical techniques were missing to assess/visualize/communicate the contributions of the interventions? In what specific ways were methods not integrated at all or incorrectly integrated into the analysis, visualization or communication?
- (e) INABILITY TO LEARN AND USE: What are some specific examples of institutions' and organizations' inability to learn and use evaluations? Were there any surprises in the inability to learn from and use evaluations? Is there any example of repeated lesson learned or recommendations that have not contributed to any substantial change or improvement?

Thursday, 4 October 2018  
11:30–13:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 41 Evaluating Complex Governance Interventions: Methodological Approaches to Analyse the Effects of Budget Support Programmes**

## Evaluating Complex Governance Interventions: Methodological Approaches to Analyse the Effects of Budget Support Programmes

*M. Orth*<sup>1</sup>

<sup>1</sup> German Institute for Development Evaluation DEval, Governance- Bi and Multilateral Development Cooperation, Bonn, Germany

Budget support is a complex aid instrument in contemporary development cooperation. At the beginning of the 2000s, it evolved into arguably the most prominent, but also most heatedly debated aid modality. Defined as a financing method to provide funding to a partner country's budget by transferring resources from an external donor to the national treasury of the partner government, it follows an extensive intervention logic and combines financial and non-financial inputs, often by multiple donors, in order to achieve multiple outcomes. There has been a long-standing methodological debate on how to assess the contribution of complex budget support programmes to the envisaged changes at different levels. Evaluation experts under the lead of the European Commission proposed a common approach to evaluating budget support, which includes a generic intervention logic. The common approach to evaluating budget support has been applied in a large number of evaluations. Despite the considerable resources invested in developing the methodological approach, fundamental difficulties in evaluating budget support remain. One of the main unsolved issues is the challenge of defining the counterfactual for budget support programmes. Moreover, the broadly used methodology – the so called "Three-Step-Approach" – is adequate to cover the large scope of budget support programmes and helps to generate a substantial body of evidence on the effectiveness of budget support, but is weak with regard to establishing clear and unambiguous causal links between the inputs and the outcomes of budget support. Even recent evaluations continue to face problems of attribution, as the explanatory power of budget support evaluations is particularly limited for the outcome and impact level. The proposed panel focuses on different methodological approaches in evaluating budget support programmes to address the above-mentioned challenges in budget support evaluations. To that purpose, the panel consists of three contributions:

- A Mechanism-Centred Approach to Evaluating Complex Aid Interventions: the Case of Accompanying Measures to General Budget Support
- Evaluating the Exit from General Budget Support: Effects of the Exit from General Budget Support
- Evaluating Budget Support Operations: A Comparison of the IEG and OECD-DAC Approaches and some Lessons

All three contributions discuss theory-based approaches to evaluating budget support, but develop and apply different theories of change to cater for the specific challenges in budget support evaluations. After short presentations of each contribution, the panel participants will discuss how and in how far the presented approaches address the challenges in budget support evaluations, ranging from the choice of counterfactual to adjustments in the theory of change to the interrelation of programme elements.

**PN 41 Evaluating Complex Governance Interventions: Methodological Approaches to Analyse the Effects of Budget Support Programmes****PC 153 - A Mechanism-Centred Approach to Evaluating Complex Aid Interventions: The Case of Accompanying Measures to General Budget Support**

*J. Schmitt*<sup>1</sup>

<sup>1</sup> *German Institute for Development Evaluation DEval, Competence Centre for Evaluation Methodologies, Bonn, Germany*

Current methodological debates related to theory-based evaluations (TBE) centre around questions on how to improve the explanatory strength of these approaches and how to integrate mechanisms as analytic concept. Particularly in complex aid interventions, when multiple elements are expected to interact and thus create an added value, exploring mechanisms as an analytical tool can be promising. This paper seeks to contribute to the discussion on the use of TBE for evaluating complex aid interventions by sharing experiences from a recent evaluation of accompanying measures to general budget support. Accompanying measures (mainly in the form of technical assistance and capacity development) are one element of the budget support package, which further encompasses financial contributions, policy dialogue, and conditionalities. We focus on interrelations between different elements of budget support and apply a mechanism-centred approach to programme theory building. After defining accompanying measures and integrating them into the intervention logic of budget support used in recent multi-donor evaluations, the paper presents key mechanisms as identified on an explorative mission to Mozambique, and shows how we validated them in an online survey, and further discussed them in expert interviews and during field research in Tanzania. For the specific example relating to two elements of budget support (policy dialogue and accompanying measures), we find that some of the hypothesized mechanisms were present and created an added value and thus increase the effectiveness of budget support as a package. The applied approach helped generate a more comprehensive implementation theory and provided insights into potential benefits and challenges of combining different elements in one programme. Beyond its use for future evaluations in the field of budget support, we argue that TBE of complex interventions can benefit from adopting such a mechanism-centred approach to create a better understanding of how different elements of the programme interact. Moreover, the focus on mechanisms when analysing programme implementation enables evaluators to improve their empirical inquiry on the identified mechanisms and to draw valid conclusions on the programme's contribution to the observed outcomes.

**PN 41 Evaluating Complex Governance Interventions: Methodological Approaches to Analyse the Effects of Budget Support Programmes****PC 154 - Evaluating the Exit from Aid: Effects of the Exit from General Budget Support**

*M. Orth*<sup>1</sup>, *G. Gotz*<sup>1</sup>

<sup>1</sup> *German Institute for Development Evaluation DEval, Governance- Bi and Multilateral Development Cooperation, Bonn, Germany*

**Rationale:** Budget support has been a favoured aid instrument when considering implementing the principles of effective aid formulated in the 2005 Paris Declaration. Evaluations find that budget support can be an effective instrument to reduce poverty. Nonetheless, budget support has increasingly come under criticism, and many bilateral donors have either partly or fully stopped using this instrument. Against this backdrop, the team reviewed existing evidence related to the effectiveness of budget support, and investigated the consequences of terminating the use of this instrument to provide recommendations that ensure resilience in the exit from aid.

**Objectives:** The evaluation aims at making lessons learned from ending and evaluating budget support available to decision makers. Specifically the evaluation aims to

- contribute to understand the consequences of ending budget support at the country level,

- identify best practices to ensure sustainability of budget support outcomes, and to
- determine approaches to mitigate potential negative effects of ending budget support.

**Brief Narrative and Justification Methodology:** This theory-based evaluation uses a comparative case study design, which is known to be comprehensive and suitable to evaluate complex interventions, but weaker regarding testing causal attributions. To address this challenge the case study design was combined with process tracing. It is an approach particularly suited to answer “if” and “how” questions and provides higher confidence in the attribution of the effect. The mixed method approach of comparative case studies and process tracing allows for a high degree of internal and external validity. For the case studies, the team conducted over 100 semi-structured interviews in Malawi, Uganda, Rwanda and Zambia and used country-specific budget analysis. In combination with an evaluation synthesis on the effects of budget support, the exit from the modality also serves as a mean to confirm that budget support programmes were responsible for the observed changes in the first place (quasi-counterfactual). Main Findings After the exit, structures for policy dialogue and coordination collapsed in all four case study countries and the fragmentation of the aid portfolio increased. Public expenditure for poverty relevant sectors declined in three of the four countries as well as reform progress in public financial management, if not bolstered by other internal/external incentives, such as pre-conditions for future loans or pressure from CSOs and media. These findings stand in stark contrast to findings on positive budget support effects prior to the exit and provide confidence that the effects of budget support were causal in the first place.

**Conclusion:** This evaluation finds that the broad and mostly unplanned exit from budget support undermined most positive effects associated with the provision of budget support. Based on these findings recommendations on exit strategies to ensure robust effects of budget support and other programmes are derived.

#### **PN 41 Evaluating Complex Governance Interventions: Methodological Approaches to Analyse the Effects of Budget Support Programmes**

### **PC 155 - Evaluating Budget Support Operations: A Comparison of the IEG and OECD-DAC Approaches and some Lessons**

*Ž. Bogetić<sup>1</sup>*

<sup>1</sup> *Independent Evaluation Group IEG, Macro-Fiscal Management, Washington DC, USA*

This paper compares the IEG and OECD-DAC approaches to budget support evaluations by focusing on the underlying methodologies and their strengths and weaknesses. It also raises questions on how the two approaches can learn from each other. Preliminary lessons are drawn for potential revisions in these approaches towards their better design and implementation. It is hoped that the paper will inform the ongoing discussions about the OECD-DAC approach to budget support evaluation in the EU as well as the IEG discussions about its own approach to evaluating budget support operations at the World Bank.

Thursday, 4 October 2018  
11:30–13:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 42 Evaluation Synthesis: Enhancing The Use Of Evaluation Findings for Decision Making And Reform Processes - Experience From IFAD And WFP**

## **Evaluation Synthesis: Enhancing the Use Of Evaluation Findings for Decision Making and Reform Processes – Experience from IFAD And WFP**

A. Cook<sup>1</sup>

<sup>1</sup> World Food Programme, Office of Evaluation, Rome, Italy

In recent years evaluation synthesis reports have gained ever greater importance in IFAD and WFP, for several reasons:

- They present evidence on thematic issues in an accessible way for a wider audience or decision makers.
- They draw from a wider range of evaluations and are therefore able to generalise.
- They are relatively time and cost-efficient and able to respond to emerging issues and requests for evidence in a timely manner.

IFAD and WFP have been experimenting with different types of synthesis methodologies and products. The session will present experiences and emerging lessons on how evaluation syntheses can be used to present evidence in a credible, yet accessible manner.

The session will start with short introduction of the session, which will be chaired by Andrea Cook (Director of Evaluation, WFP). This will be followed by presentations from IFAD and WFP on different types of synthesis methods and products and a discussion exploring how synthesis reports can be effectively used to inform decision making and reform processes.

Andrea Cook is the Director of Evaluation in WFP. Before joining WFP, she was Director of Evaluation at UNFPA and worked for over twenty years in international development, in the UK, Asia, Africa, the Caribbean and Eastern Europe in a variety of roles spanning policy, programme, management and evaluation.

Johanna Pennarz (IFAD Lead Evaluation Officer) will reflect on how IOE has moved to a more rigorous approach to evaluation synthesis to address the increasing demand for independent and focussed analysis of selected topics by the IFAD board. More recently the adoption of systematic review methodologies has been supported by the use of qualitative software (NVIVO) which has enabled IFAD to draw evidence from a larger number of evaluation reports. The panellist will provide a succinct presentation on this methodology (and the application of NVIVO as part of this) using the example of the recent evaluation synthesis on inclusive rural finance.

Hansdeep Khaira (IFAD Evaluation Officer): will present the methodology for synthesising results from different impact evaluations within the context of country strategy and programme evaluations, using the example of a recent evaluation in Kenya. The point of departure of the methodology is identifying commonalities within the theory of change, which can be focussed on a selected theme, e.g. value chains. Synthesising the results of impact evaluations (that rely on mixed methods) enables the evaluation to infer what has worked, or not, and under what conditions. This involves focussing on individual elements of the theme (along the chain) and assessing their role in driving the impact on selected indicators of rural poverty.

Deborah McWhinney (WFP Senior Evaluation Officer) will share lessons from the recent experience at WFP including: synthesising the results of a series of 58 Operations Evaluations in humanitarian and development contexts; synthesising the results of four series of impact evaluations; and the development and testing of Atlas.ti software to better support WFP to access evidence from an increasing number of centralised and decentralised evaluation reports to inform country strategy, policy making and corporate decision making.

**PN 42 Evaluation Synthesis: Enhancing The Use Of Evaluation Findings for Decision Making  
And Reform Processes - Experience From IFAD And WFP****PC 286 - Panel Contribution**

A. Cook<sup>1</sup>

<sup>1</sup> World Food Programme, Office of Evaluation, Rome, Italy

Andrea Cook (Director of Evaluation WFP) will chair a panel discussion to explore how synthesis reports can be effectively used to inform decision making and reform process. This will include lessons on conceptual and methodological issues arising from experience in IFAD and WFP to date.

**PN 42 Evaluation Synthesis: Enhancing The Use Of Evaluation Findings for Decision Making  
And Reform Processes - Experience From IFAD And WFP****PC 156 - Panel Contribution**

J. Pennarz<sup>1</sup>

<sup>1</sup> IFAD, Independent Office of Evaluation, Rome, Italy

Johanna Pennarz (IFAD Lead Evaluation Officer) will reflect on how IOE has moved to a more rigorous approach to evaluation synthesis to address the increasing demand for independent and focussed analysis of selected topics by the IFAD board. More recently the adoption of systematic review methodologies has been supported by the use of qualitative software (NVIVO) which has enabled IFAD to draw evidence from a larger number of evaluation reports. The panellist will provide a succinct presentation on this methodology (and the application of NVIVO as part of this) using the example of the recent evaluation synthesis on inclusive rural finance.

**PN 42 Evaluation Synthesis: Enhancing The Use Of Evaluation Findings for Decision Making  
And Reform Processes - Experience From IFAD And WFP****PC 157 - Panel Contribution**

H. Khaira<sup>1</sup>

<sup>1</sup> IFAD, Independent Office of Evaluation, Rome, Italy

Hansdeep Khaira (IFAD Evaluation Officer): will present the methodology for synthesising results from different impact evaluations within the context of country strategy and programme evaluations, using the example of a recent evaluation in Kenya. The point of departure of the methodology is identifying commonalities within the theory of change, which can be focussed on a selected theme, e.g. value chains. Synthesising the results of impact evaluations (that rely on mixed methods) enables the evaluation to infer what has worked, or not, and under what conditions. This involves focussing on individual elements of the theme (along the chain) and assessing their role in driving the impact on selected indicators of rural poverty.

**PN 42 Evaluation Synthesis: Enhancing The Use Of Evaluation Findings for Decision Making  
And Reform Processes - Experience From IFAD And WFP**

**PC 158 - Panel Contribution**

*D. McWhinney*<sup>1</sup>

<sup>1</sup> *World Food Programme, Office of Evaluation, Rome, Italy*

Deborah McWhinney (WFP Senior Evaluation Officer ) will share lessons from the recent experience at WFP including: synthesising the results of a series of 58 Operations Evaluations in humanitarian and development contexts; synthesising the results of four series of impact evaluations; and the development and testing of Atlas.ti software to better support WFP to access evidence from an increasing number of centralised and decentralised evaluation reports to inform country strategy, policy making and corporate decision making.

Thursday, 4 October 2018  
11:30–13:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 43 Meeting of Thematic Working Group for Emerging Evaluators**

## **Meeting of Thematic Working Group for Emerging Evaluators**

M. Branco<sup>1</sup>

<sup>1</sup> *Independent consultant, Porto, Portugal*

During this session, we will present the work of the EES Thematic Working Group for Emerging Evaluators and reflect together on how the group can be further developed in the future. Anyone who is interested is warmly invited to participate and welcome to contribute fresh ideas.

Thursday, 4 October 2018  
11:30–13:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 44 Designing and Implementing Adaptive Evaluations for Complex Programmes in Conflict Zone**

## **Designing and Implementing Adaptive Evaluations for Complex Programmes in Conflict Zones**

*S. Lemire<sup>1</sup>, G. Freer<sup>2</sup>, J. Thakrar<sup>3</sup>, K. Reid<sup>4</sup>*

<sup>1</sup> University Of California- Los Angeles, Department of Education and Information Studies, Los Angeles, USA

<sup>2</sup> Insight Strategies, Not applicable, Johannesburg, South Africa

<sup>3</sup> WYG Associates, Not applicable, Cambridgeshire, United Kingdom

<sup>4</sup> WYG, Not applicable, Edinburgh, United Kingdom

Increasingly complex and dynamic development programmes are calling for adaptive and multi-method evaluations. Adaptive and responsive designs are particularly salient for evaluations in high conflict zones or fragile environments, where the programmes being evaluated must often adapt in rapid response to unpredictable disruptions. Grounded on real-world examples, this panel illustrates and considers the extent to which and how evaluations grounded on theory-based, developmental, and quasi-experimental design components can be adapted to counter the risks of doing evaluations in fragile environments and contexts of conflict. Speaking directly to the conference theme of evaluation in turbulent times, the panel will share their experiences and lessons learned in the design and implementation of evaluations in these contexts, identifying some of their successes and weaknesses as well as future challenges they have yet to overcome. The panellists will discuss their application of specific designs and approaches for adaptive evaluations that maintain methodological integrity and rigour, while still allowing for evolving programmes and shifting contexts. The illustrated cases comprising the panel stem from three independent evaluations of DFID-funded programmes. The featured cases are:

1. The combination of a theory-based evaluation approach and a matching design applied in a multi-year evaluation of an adaptive market development programme (Propcom Mai-karfi) in Northern Nigeria;
2. The use of developmental evaluation as a framework for real time learning in an adaptive evaluation of the Zambia Accountability Programme (ZAP), which seeks to improve accountability and responsiveness in the delivery of public goods and services in Zambia;
3. The application of contribution analysis and its use of mixed-method data to create discussion and learning in a multi-year evaluation of a market development programme, the Comprehensive Agricultural and Rural Development Facility (CARD-F), in rural Afghanistan.

Discussions will highlight how the different approaches and designs have been employed and reflect on their strengths and limitations, as compared with other solutions available. We aim to generate debate and audience interaction through discussion of past experiences and ideas, such as the importance of:

- Evaluability and multi-faceted environments – choosing the right tools for the right job in the right environment;
- The application of mixed-methods;
- The impact of successful working relationships with partners;
- Adjusting approaches for ongoing learning and adaptation.

The panel will aim to discuss the successes and difficulties of conducting evaluations in these environments; the feasibility of methods used and lessons learned. Each panelist will provide a short presentation of their evaluation case lasting no longer than 15 minutes. The three panel presentations will be followed by an open discussion with the audience on challenges and lessons learned with adaptive programmes in conflict zones and fragile contexts.

**PN 44 Designing and Implementing Adaptive Evaluations for Complex Programmes in Conflict Zone**

**PC 159 - Adapting to Adaptive Programming: Rethinking Roles, Relations and Relevance of Evaluation In a Complex Accountability Programme**

J. Thakrar<sup>1</sup>

<sup>1</sup> WYG Associates, Not applicable, Cambridgeshire, United Kingdom

Jayshree Thakrar will discuss the use of developmental evaluation as a framework for real time learning in an adaptive evaluation of the Zambia Accountability Programme (ZAP), which seeks to improve accountability and responsiveness in the delivery of public goods and services in Zambia.

**PN 44 Designing and Implementing Adaptive Evaluations for Complex Programmes in Conflict Zone**

**PC 160 - Contribution Analysis and Conflict Zones: Assessing Impact in Afghanistan**

K. Reid<sup>1</sup>

<sup>1</sup> WYG, Not applicable, Edinburgh, United Kingdom

Kyla Reid will share lessons learned from an application of contribution analysis and its use of mixed-method data to create discussion and learning in a multi-year evaluation of a market development programme, the Comprehensive Agricultural and Rural Development Facility (CARD-F), in rural Afghanistan.

**PN 44 Designing and Implementing Adaptive Evaluations for Complex Programmes in Conflict Zone**

**PC 161 - Measuring Moving Targets: Evaluating Adaptive Market Development Programmes in Fragile Contexts**

G. Freer<sup>1</sup>

<sup>1</sup> Insight Strategies, Not applicable, Johannesburg, South Africa

Gordon Freer will discuss the combination of a theory-based evaluation approach and a matching design applied in a multi-year evaluation of an adaptive market development programme (Propcom Mai-karfi) in Northern Nigeria.

Thursday, 4 October 2018  
15:00–16:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### PN 45 Evaluation Standards and Competencies for Resilient Societies

## Evaluation Standards and Competencies for Resilient Societies

*D. Schroeter<sup>1</sup>, B. Watts<sup>2</sup>, J. Hense<sup>3</sup>, T. Widmer<sup>4</sup>, S. Leahy<sup>5</sup>, A. Cook<sup>6</sup>, J. Flentge<sup>7</sup>, M. Tarsilla<sup>8</sup>*

<sup>1</sup> Western Michigan University, School of Public Affairs and Administration, Kalamazoo, USA

<sup>2</sup> Western Michigan University, Evaluation Center, Kalamazoo, USA

<sup>3</sup> Justus-Liebig-Universität Gießen, Hochschuldidaktik und Evaluation, Giessen, Germany

<sup>4</sup> University of Zurich, Department of Political Science, Zurich, Switzerland

<sup>5</sup> ARTD Consultants, Australasian Evaluation Society, Sydney, Australia

<sup>6</sup> UN World Food Programme, UN Evaluation Group, Rome, Italy

<sup>7</sup> UN World Food Programme, Officer of Evaluation, Rome, Italy

<sup>8</sup> UNICEF, West and Central Africa Regional Office, Yoff Dakar, Senegal

The purpose of this panel is to engage presenters and the audience in a critical discussion on two key topics that appear relevant to the resilience of the evaluation profession in the future, both within and outside of Europe. First, what regional and global evaluation standards and competencies have been developed so far and how? Second, what type of relationship exists between the existing evaluation standards and the evaluation competency frameworks currently available? The panelists, who have over 80 years of combined evaluation experience, will discuss their direct experience in developing and using evaluation standards and competencies in numerous countries and organizations. Regardless of their distinct perspectives on these issues, panelists will share their experience and expertise in relation to the following questions:

How have they contributed to the development of evaluation standard?

How have the developed standards been adapted over time to the needs of the different evaluation communities which they were supposed to serve, both regionally and globally?

- How are evaluation standards used, and by whom?
- What is the relationship between evaluation standards and evaluation competency frameworks?
- How have evaluation standards been updated to reflect the recent developments in theory and practice (including emerging barriers and opportunities)?
- Which actors are the most engaged in standards and competency development?
- What is the role of VOPEs, universities, and other stakeholders in the development of evaluation standards?
- To what extent and how are standards and competencies independent from, and relevant and responsive to the needs of their users?

The session chair will aggregate key comparative documentation from the participating (and other) organizations to set the stage for facilitating debate and discussion between the panelists and audience. The panel is equally relevant to novice and more seasoned evaluation practitioners, managers, commissioners and users as it raises their awareness about the opportunities and intricacies of developing evaluation standards and competencies worldwide. Furthermore, the panel aims to clarify (through the use of case studies) the extent to which evaluation standards and competency frameworks are furthering the utility, feasibility, propriety, and accuracy of evaluation endeavors in a number of organizational and cultural settings. In doing so, the panel is intended to advance the current discourse on how to tap into the opportunities and address the challenges associated with the innovation of the existing evaluation standards. Lastly, the panel is expected to provide the audience with a number of concrete scenarios in which evaluation professionals "resilience" was tested and amply accrued in the past. In an attempt to enhance the audience capacity to deal with the currently tumultuous political climate characterized by "fake news" and "alternative facts," the panel will demonstrate how critical discussions across cultures and openness to "otherness" not only foster adaptiveness in evaluation practice but also promote more effective action towards increasingly resilient and equitable societies.

**PN 45 Evaluation Standards and Competencies for Resilient Societies****PC 162 - The Role and Future of Standards in North America**

D. Schroeter<sup>1</sup>, B. Watts<sup>2</sup>, J. Hense<sup>3</sup>, T. Widmer<sup>4</sup>

<sup>1</sup> Western Michigan University, School of Public Affairs and Administration, Kalamazoo, USA

<sup>2</sup> Western Michigan University, Evaluation Center, Kalamazoo, USA

<sup>3</sup> Justus-Liebig-Universität Gießen, Hochschuldidaktik und Evaluation, Giessen, Germany

<sup>4</sup> University of Zurich, Department of Political Science, Zurich, Switzerland

The purpose of this panel is to engage presenters and the audience in a critical discussion on two key topics that appear relevant to the resilience of the evaluation profession in the future, both within and outside of Europe. First, what regional and global evaluation standards and competencies have been developed so far and how? Second, what type of relationship exists between the existing evaluation standards and the evaluation competency frameworks currently available?

The panelists, who have over 80 years of combined evaluation experience, will discuss their direct experience in developing and using evaluation standards and competencies in numerous countries and organizations. Regardless of their distinct perspectives on these issues, panelists will share their experience and expertise in relation to the following questions:

- How have they contributed to the development of evaluation standard?
- How have the developed standards been adapted over time to the needs of the different evaluation communities which they were supposed to serve, both regionally and globally?
- How are evaluation standards used, and by whom?
- What is the relationship between evaluation standards and evaluation competency frameworks?
- How have evaluation standards been updated to reflect the recent developments in theory and practice (including emerging barriers and opportunities)?
- Which actors are the most engaged in standards and competency development?
- What is the role of VOPEs, universities, and other stakeholders in the development of evaluation standards?
- To what extent and how are standards and competencies independent from, and relevant and responsive to the needs of their users?

The session chair will aggregate key comparative documentation from the participating (and other) organizations to set the stage for facilitating debate and discussion between the panelists and audience.

The panel is equally relevant to novice and more seasoned evaluation practitioners, managers, commissioners and users as it raises their awareness about the opportunities and intricacies of developing evaluation standards and competencies worldwide. Furthermore, the panel aims to clarify (through the use of case studies) the extent to which evaluation standards and competency frameworks are furthering the utility, feasibility, propriety, and accuracy of evaluation endeavors in a number of organizational and cultural settings. In doing so, the panel is intended to advance the current discourse on how to tap into the opportunities and address the challenges associated with the innovation of the existing evaluation standards. Lastly, the panel is expected to provide the audience with a number of concrete scenarios in which evaluation professionals "resilience" was tested and amply accrued in the past. In an attempt to enhance the audience capacity to deal with the currently tumultuous political climate characterized by "fake news" and "alternative facts," the panel will demonstrate how critical discussions across cultures and openness to "otherness" not only foster adaptiveness in evaluation practice but also promote more effective action towards increasingly resilient and equitable societies.

(The session chair will provide an overview and raise questions for discussion, debate, and exchange between the panelists. Brad will provide insight into the North American standards discussion. It will be the chair's responsibility to maximize contributions and exchange between all panelists.)

**PN 45 Evaluation Standards and Competencies for Resilient Societies****PC 163 - Evaluation Standards and Competencies in Germany and Austria**

*D. Schroeter<sup>1</sup>, B. Watts<sup>2</sup>, J. Hense<sup>3</sup>, T. Widmer<sup>4</sup>*

<sup>1</sup> *Western Michigan University, School of Public Affairs and Administration, Kalamazoo, USA*

<sup>2</sup> *Western Michigan University, Evaluation Center, Kalamazoo, USA*

<sup>3</sup> *Justus-Liebig-Universität Gießen, Department of Psychology, Giessen, Germany*

<sup>4</sup> *University of Zurich, Department of Political Science, Zurich, Switzerland*

The purpose of this panel is to engage presenters and the audience in a critical discussion on two key topics that appear relevant to the resilience of the evaluation profession in the future, both within and outside of Europe. First, what regional and global evaluation standards and competencies have been developed so far and how? Second, what type of relationship exists between the existing evaluation standards and the evaluation competency frameworks currently available?

The panelists, who have over 80 years of combined evaluation experience, will discuss their direct experience in developing and using evaluation standards and competencies in numerous countries and organizations. Regardless of their distinct perspectives on these issues, panelists will share their experience and expertise in relation to the following questions:

- How have they contributed to the development of evaluation standard?
- How have the developed standards been adapted over time to the needs of the different evaluation communities which they were supposed to serve, both regionally and globally?
- How are evaluation standards used, and by whom?
- What is the relationship between evaluation standards and evaluation competency frameworks?
- How have evaluation standards been updated to reflect the recent developments in theory and practice (including emerging barriers and opportunities)?
- Which actors are the most engaged in standards and competency development?
- What is the role of VOPEs, universities, and other stakeholders in the development of evaluation standards?
- To what extent and how are standards and competencies independent from, and relevant and responsive to the needs of their users?

The session chair will aggregate key comparative documentation from the participating (and other) organizations to set the stage for facilitating debate and discussion between the panelists and audience.

The panel is equally relevant to novice and more seasoned evaluation practitioners, managers, commissioners and users as it raises their awareness about the opportunities and intricacies of developing evaluation standards and competencies worldwide. Furthermore, the panel aims to clarify (through the use of case studies) the extent to which evaluation standards and competency frameworks are furthering the utility, feasibility, propriety, and accuracy of evaluation endeavors in a number of organizational and cultural settings. In doing so, the panel is intended to advance the current discourse on how to tap into the opportunities and address the challenges associated with the innovation of the existing evaluation standards. Lastly, the panel is expected to provide the audience with a number of concrete scenarios in which evaluation professionals "resilience" was tested and amply accrued in the past. In an attempt to enhance the audience capacity to deal with the currently tumultuous political climate characterized by "fake news" and "alternative facts," the panel will demonstrate how critical discussions across cultures and openness to "otherness" not only foster adaptiveness in evaluation practice but also promote more effective action towards increasingly resilient and equitable societies.

The session chair will provide an overview and raise questions for discussion, debate, and exchange between the panelists. Professor Hense will provide particular insight into the German language discourse about standards and competencies. It will be the chair's responsibility to maximize contributions and exchange between all panelists.

**PN 45 Evaluation Standards and Competencies for Resilient Societies****PC 164 - The Swiss Experience with Evaluation Standards**

*D. Schroeter<sup>1</sup>, B. Watts<sup>2</sup>, J. Hense<sup>3</sup>, T. Widmer<sup>4</sup>*

<sup>1</sup> Western Michigan University, School of Public Affairs and Administration, Kalamazoo, USA

<sup>2</sup> Western Michigan University, Evaluation Center, Kalamazoo, USA

<sup>3</sup> Justus-Liebig-Universität Gießen, Hochschuldidaktik und Evaluation, Giessen, Germany

<sup>4</sup> University of Zurich, Department of Political Science, Zurich, Switzerland

The purpose of this panel is to engage presenters and the audience in a critical discussion on two key topics that appear relevant to the resilience of the evaluation profession in the future, both within and outside of Europe. First, what regional and global evaluation standards and competencies have been developed so far and how? Second, what type of relationship exists between the existing evaluation standards and the evaluation competency frameworks currently available?

The panelists, who have over 80 years of combined evaluation experience, will discuss their direct experience in developing and using evaluation standards and competencies in numerous countries and organizations. Regardless of their distinct perspectives on these issues, panelists will share their experience and expertise in relation to the following questions:

- How have they contributed to the development of evaluation standard?
- How have the developed standards been adapted over time to the needs of the different evaluation communities which they were supposed to serve, both regionally and globally?
- How are evaluation standards used, and by whom?
- What is the relationship between evaluation standards and evaluation competency frameworks?
- How have evaluation standards been updated to reflect the recent developments in theory and practice (including emerging barriers and opportunities)?
- Which actors are the most engaged in standards and competency development?
- What is the role of VOPEs, universities, and other stakeholders in the development of evaluation standards?
- To what extent and how are standards and competencies independent from, and relevant and responsive to the needs of their users?

The session chair will aggregate key comparative documentation from the participating (and other) organizations to set the stage for facilitating debate and discussion between the panelists and audience.

The panel is equally relevant to novice and more seasoned evaluation practitioners, managers, commissioners and users as it raises their awareness about the opportunities and intricacies of developing evaluation standards and competencies worldwide. Furthermore, the panel aims to clarify (through the use of case studies) the extent to which evaluation standards and competency frameworks are furthering the utility, feasibility, propriety, and accuracy of evaluation endeavors in a number of organizational and cultural settings. In doing so, the panel is intended to advance the current discourse on how to tap into the opportunities and address the challenges associated with the innovation of the existing evaluation standards. Lastly, the panel is expected to provide the audience with a number of concrete scenarios in which evaluation professionals "resilience" was tested and amply accrued in the past. In an attempt to enhance the audience capacity to deal with the currently tumultuous political climate characterized by "fake news" and "alternative facts," the panel will demonstrate how critical discussions across cultures and openness to "otherness" not only foster adaptiveness in evaluation practice but also promote more effective action towards increasingly resilient and equitable societies.

The session chair will provide an overview and raise questions for discussion, debate, and exchange between the panelists. Dr. Widmer will highlight the Swiss experience with evaluation standards. It will be the chair's responsibility to maximize contributions and exchange between all panelists.

**PN 45 Evaluation Standards and Competencies for Resilient Societies****PC 165 - Evaluation Standards and Competencies:  
An Australasian Perspective**

*D. Schroeter<sup>1</sup>, B. Watts<sup>2</sup>, J. Hense<sup>3</sup>, T. Widmer<sup>4</sup>, S. Leahy<sup>5</sup>*

<sup>1</sup> *Western Michigan University, School of Public Affairs and Administration, Kalamazoo, USA*

<sup>2</sup> *Western Michigan University, Evaluation Center, Kalamazoo, USA*

<sup>3</sup> *Justus-Liebig-Universität Gießen, Hochschuldidaktik und Evaluation, Giessen, Germany*

<sup>4</sup> *University of Zurich, Department of Political Science, Zurich, Switzerland*

<sup>5</sup> *ARTD Consultants, Australasian Evaluation Society, Sydney, Australia*

The purpose of this panel is to engage presenters and the audience in a critical discussion on two key topics that appear relevant to the resilience of the evaluation profession in the future, both within and outside of Europe. First, what regional and global evaluation standards and competencies have been developed so far and how? Second, what type of relationship exists between the existing evaluation standards and the evaluation competency frameworks currently available?

The panelists, who have over 80 years of combined evaluation experience, will discuss their direct experience in developing and using evaluation standards and competencies in numerous countries and organizations. Regardless of their distinct perspectives on these issues, panelists will share their experience and expertise in relation to the following questions:

- How have they contributed to the development of evaluation standard?
- How have the developed standards been adapted over time to the needs of the different evaluation communities which they were supposed to serve, both regionally and globally?
- How are evaluation standards used, and by whom?
- What is the relationship between evaluation standards and evaluation competency frameworks?
- How have evaluation standards been updated to reflect the recent developments in theory and practice (including emerging barriers and opportunities)?
- Which actors are the most engaged in standards and competency development?
- What is the role of VOPEs, universities, and other stakeholders in the development of evaluation standards?
- To what extent and how are standards and competencies independent from, and relevant and responsive to the needs of their users?

The session chair will aggregate key comparative documentation from the participating (and other) organizations to set the stage for facilitating debate and discussion between the panelists and audience.

The panel is equally relevant to novice and more seasoned evaluation practitioners, managers, commissioners and users as it raises their awareness about the opportunities and intricacies of developing evaluation standards and competencies worldwide. Furthermore, the panel aims to clarify (through the use of case studies) the extent to which evaluation standards and competency frameworks are furthering the utility, feasibility, propriety, and accuracy of evaluation endeavors in a number of organizational and cultural settings. In doing so, the panel is intended to advance the current discourse on how to tap into the opportunities and address the challenges associated with the innovation of the existing evaluation standards. Lastly, the panel is expected to provide the audience with a number of concrete scenarios in which evaluation professionals "resilience" was tested and amply accrued in the past. In an attempt to enhance the audience capacity to deal with the currently tumultuous political climate characterized by "fake news" and "alternative facts," the panel will demonstrate how critical discussions across cultures and openness to "otherness" not only foster adaptiveness in evaluation practice but also promote more effective action towards increasingly resilient and equitable societies.

The session chair will provide an overview and raise questions for discussion, debate, and exchange between the panelists. Sue will highlight the Australasian experience with evaluation standards and competencies. It will be the chair's responsibility to maximize contributions and exchange between all panelists.

**PN 45 Evaluation Standards and Competencies for Resilient Societies****PC 166 - The Perspective of the UNEG**

*D. Schroeter<sup>1</sup>, B. Watts<sup>2</sup>, J. Hense<sup>3</sup>, T. Widmer<sup>4</sup>, A. Cook<sup>5</sup>*

<sup>1</sup> *Western Michigan University, School of Public Affairs and Administration, Kalamazoo, USA*

<sup>2</sup> *Western Michigan University, Evaluation Center, Kalamazoo, USA*

<sup>3</sup> *Justus-Liebig-Universität Gießen, Hochschuldidaktik und Evaluation, Giessen, Germany*

<sup>4</sup> *University of Zurich, Department of Political Science, Zurich, Switzerland*

<sup>5</sup> *Departing Vice Chair UNEG, Director of Evaluation- UN World Food Programme, Rome, Italy*

The purpose of this panel is to engage presenters and the audience in a critical discussion on two key topics that appear relevant to the resilience of the evaluation profession in the future, both within and outside of Europe. First, what regional and global evaluation standards and competencies have been developed so far and how? Second, what type of relationship exists between the existing evaluation standards and the evaluation competency frameworks currently available?

The panelists, who have over 80 years of combined evaluation experience, will discuss their direct experience in developing and using evaluation standards and competencies in numerous countries and organizations. Regardless of their distinct perspectives on these issues, panelists will share their experience and expertise in relation to the following questions:

- How have they contributed to the development of evaluation standard?
- How have the developed standards been adapted over time to the needs of the different evaluation communities which they were supposed to serve, both regionally and globally?
- How are evaluation standards used, and by whom?
- What is the relationship between evaluation standards and evaluation competency frameworks?
- How have evaluation standards been updated to reflect the recent developments in theory and practice (including emerging barriers and opportunities)?
- Which actors are the most engaged in standards and competency development?
- What is the role of VOPEs, universities, and other stakeholders in the development of evaluation standards?
- To what extent and how are standards and competencies independent from, and relevant and responsive to the needs of their users?

The session chair will aggregate key comparative documentation from the participating (and other) organizations to set the stage for facilitating debate and discussion between the panelists and audience.

The panel is equally relevant to novice and more seasoned evaluation practitioners, managers, commissioners and users as it raises their awareness about the opportunities and intricacies of developing evaluation standards and competencies worldwide. Furthermore, the panel aims to clarify (through the use of case studies) the extent to which evaluation standards and competency frameworks are furthering the utility, feasibility, propriety, and accuracy of evaluation endeavors in a number of organizational and cultural settings. In doing so, the panel is intended to advance the current discourse on how to tap into the opportunities and address the challenges associated with the innovation of the existing evaluation standards. Lastly, the panel is expected to provide the audience with a number of concrete scenarios in which evaluation professionals "resilience" was tested and amply accrued in the past. In an attempt to enhance the audience capacity to deal with the currently tumultuous political climate characterized by "fake news" and "alternative facts," the panel will demonstrate how critical discussions across cultures and openness to "otherness" not only foster adaptiveness in evaluation practice but also promote more effective action towards increasingly resilient and equitable societies.

The session chair will provide an overview and raise questions for exchange between the panelists. Andrea will highlight her experience with the formulation of the UNEG's "Norms and Standards" and "Evaluation Competency Framework." It will be the chair's responsibility to maximize contributions and exchange between all panelists.

**PN 45 Evaluation Standards and Competencies for Resilient Societies****PC 167 - Competency Development and Standards in WFP**

D. Schroeter<sup>1</sup>, B. Watts<sup>2</sup>, J. Hense<sup>3</sup>, T. Widmer<sup>4</sup>, J. Flentge<sup>5</sup>

<sup>1</sup> Western Michigan University, School of Public Affairs and Administration, Kalamazoo, USA

<sup>2</sup> Western Michigan University, Evaluation Center, Kalamazoo, USA

<sup>3</sup> Justus-Liebig-Universität Gießen, Hochschuldidaktik und Evaluation, Giessen, Germany

<sup>4</sup> University of Zurich, Department of Political Science, Zurich, Switzerland

<sup>5</sup> UN World Food Programme, Officer of Evaluation, Rome, Italy

The purpose of this panel is to engage presenters and the audience in a critical discussion on two key topics that appear relevant to the resilience of the evaluation profession in the future, both within and outside of Europe. First, what regional and global evaluation standards and competencies have been developed so far and how? Second, what type of relationship exists between the existing evaluation standards and the evaluation competency frameworks currently available?

The panelists, who have over 80 years of combined evaluation experience, will discuss their direct experience in developing and using evaluation standards and competencies in numerous countries and organizations. Regardless of their distinct perspectives on these issues, panelists will share their experience and expertise in relation to the following questions:

- How have they contributed to the development of evaluation standard?
- How have the developed standards been adapted over time to the needs of the different evaluation communities which they were supposed to serve, both regionally and globally?
- How are evaluation standards used, and by whom?
- What is the relationship between evaluation standards and evaluation competency frameworks?
- How have evaluation standards been updated to reflect the recent developments in theory and practice (including emerging barriers and opportunities)?
- Which actors are the most engaged in standards and competency development?
- What is the role of VOPEs, universities, and other stakeholders in the development of evaluation standards?
- To what extent and how are standards and competencies independent from, and relevant and responsive to the needs of their users?

The session chair will aggregate key comparative documentation from the participating (and other) organizations to set the stage for facilitating debate and discussion between the panelists and audience.

The panel is equally relevant to novice and more seasoned evaluation practitioners, managers, commissioners and users as it raises their awareness about the opportunities and intricacies of developing evaluation standards and competencies worldwide. Furthermore, the panel aims to clarify (through the use of case studies) the extent to which evaluation standards and competency frameworks are furthering the utility, feasibility, propriety, and accuracy of evaluation endeavors in a number of organizational and cultural settings. In doing so, the panel is intended to advance the current discourse on how to tap into the opportunities and address the challenges associated with the innovation of the existing evaluation standards. Lastly, the panel is expected to provide the audience with a number of concrete scenarios in which evaluation professionals "resilience" was tested and amply accrued in the past. In an attempt to enhance the audience capacity to deal with the currently tumultuous political climate characterized by "fake news" and "alternative facts," the panel will demonstrate how critical discussions across cultures and openness to "otherness" not only foster adaptiveness in evaluation practice but also promote more effective action towards increasingly resilient and equitable societies.

The session chair will provide an overview and raise questions for discussion, debate, and exchange between the panelists. Jacqueline will highlight her experience with competencies and standards in the World Food Programme. It will be the chair's responsibility to maximize contributions and exchange between all panelists.

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 46 Rethinking Evaluation Approaches and Skills for Development and Human Rights Programmes in Context of Fragility, Protracted Conflict and Structural Violence**

## Rethinking Evaluation Approaches and Skills for Development and Human Rights Programmes in Context of Fragility, Protracted Conflict and Structural Violence

*E. Rotondo*<sup>1</sup>

<sup>1</sup> International consultant, Lima, Peru

**Rationale:** The World Bank estimates that there are now two billion people who live in countries where development outcomes are affected by fragility, conflict, and violence. With 80% of all humanitarian needs now being driven by conflict and the increased tension between the robustness of evidence and duty of care, it is the need of the time to look into current state of play in evaluating aid in conflict and share best ways of working in such contexts. World wide monitoring and evaluation agenda have focused mainly on the social policy issues, and, to a lesser extent, on limit situations such as crises, migration, violence or conflicts. This opens the discussion towards evaluations with differentiated and relevant approaches for such limit situations. Likewise, the demand and need for effective protection systems and guarantee of effective rights that point towards the prevention, elimination and non-repetition of these limit situations are increasing. Added to this, there are few spaces in the evaluation communities to share technical aspects such as the approaches used, the methodologies or the specific competences (or capacities) that the evaluators must have to approach these exercises, as well as their complementarities, complexities and challenges.

### Objectives sought:

- To share lessons learned, positive and negative, in the design and development of evaluations in fragile, volatile and violent contexts.
- Collect evaluation practices of programs oriented to personal and community resilience, developed in various continents, to propose good practices with implications for gender equity and human rights approaches.
- Suggest basic skills necessary for evaluators training to evaluate programs oriented to social and personal resilience in fragile and violent contexts.
- Brief narrative and justification

In several continents, evaluations have been carried out in fragile and violent contexts, obtaining lessons learned at a conceptual and methodological level, including reflections on the competencies needed in the training of evaluators to undertake evaluations in these contexts.

The basic questions that the panel will seek to answer will be:

- How to evaluate in fragile and volatile contexts and how to build theories of change?
- What does it mean to assess resilience, what kind of indicators should be used?
- How to incorporate the perspective of different actors, especially those most affected, in the assessment of the changes to their lives?
- How to collect expected and unintended results?
- What basic skills do evaluators need for this kind of environment?
- How to identify gender gaps?

The panel will present lessons learned and best practices that attempt to answer these questions from evaluative practices carried out in Asia, Latin America and other continents.

**PN 46 Rethinking Evaluation Approaches and Skills for Development and Human Rights Programmes in Context of Fragility, Protracted Conflict and Structural Violence****PC 169 - Challenges to Incorporating Gender When Evaluating Conflict**

A. Arbulu<sup>1</sup>

<sup>1</sup> *International Consultant, Consultant, Panama*

The focus of her presentation will be looking at the challenges faced by evaluators in conflict and volatile environments when trying to ensure that gender is treated as an equally urgent priority to other rights. Skills necessary include careful design as well as the need for interview methodologies that allow biases to surface in a non-threatening manner and participatory processes that allow for discussion and identification of solutions by the stakeholders as a way to ensure that the recommendations are relevant and feasible.

**PN 46 Rethinking Evaluation Approaches and Skills for Development and Human Rights Programmes in Context of Fragility, Protracted Conflict and Structural Violence****PC 290 - Seven Principles of Evaluating in contexts of Fragility, Conflict and Violence**

H. Hassnain<sup>1</sup>

<sup>1</sup> *International Development Evaluation Association - IDEAS, Board Member- Chair of Evaluation in Fragility-Conflict and Violence Thematic Interest Group and Founder of Pakistan Evaluation Association, London, United Kingdom*

In this presentation he will shed some light on the best practices and challenges faced in conducting evaluations in situations that fluid, complex and volatile. There will be real time examples and case studies from his 15 years of work in some of the worst conflict-affected and fragile countries in the world as well as some plans from the IDEAS' Thematic Interest Group (TIG) 'Evaluation in Fragility Conflict and Violence (EvalFCV)'. Hur will also, share examples of some proven and innovative evaluation approaches used in highly unpredictable, fluid, complex and volatile environments. This includes the use of story-based mobile data collection tool Sprockler and Outcome Harvesting in rural Pakistan, child-led studies in Afghanistan, Jordan, DRC and Uganda, youth Volunteer led evaluations in Sierra Leone and some of the real time challenges from continuously changing contexts of South Sudan.

**PN 46 Rethinking Evaluation Approaches and Skills for Development and Human Rights Programmes in Context of Fragility, Protracted Conflict and Structural Violence****PC 170 - Methodological Challenges and Evaluators Competences to Evaluate Resilience Oriented Programs. Lessons Learned from De Latinamerican Región**

E. Rotondo<sup>1</sup>

<sup>1</sup> *International consultant, Lima, Peru*

The presentation will address the conceptual and methodological evaluation lessons learned of resilient oriented programs in changing contexts of conflict and social violence that affect women, children and adolescents in particular. The presentation will present a proposal of competencies necessary for internal and external evaluators to facilitate processes, learn and improve evaluation practice.

Thursday, 4 October 2018  
15:00–16:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 47 A Framework for Accelerating Access to Science, Technology and Evaluation Results (FAASTER) for Global Health Impact**

## **A Framework for Accelerating Access to Science, Technology and Evaluation Results (FAASTER) for Global Health Impact**

*A. Djapovic*<sup>1</sup>

<sup>1</sup> *United States Agency for International Development, Bureau for Global Health, Arlington, USA*

The timeline from technology development or research findings to delivery for public health impact has been notably slow. Our Framework for Accelerating Access to Science, Technology and Evaluation Results (FAASTER) and associated tools was developed to enhance translation and utilization of innovations into PEPFAR programs. All PEPFAR implementing agencies participated in developing the standardized approach to culling and classifying research and evaluation (RE) activities. An inventory was developed of PEPFAR-supported RE activities funded between 2009 and 2016 that included defining the 21 data-elements used to classify activities, data quality and validation, and annotating governance structures and permissions. User friendly tools were developed to sort, summarize and visualize the inventory data by geography, program area, research methodology, sub-population, as well as key information for initiating data access. Geospatial presentation was done using ArcGIS. Program monitoring and evaluation data were not included.

The process uncovered that without an overall framework and a tracking system, many of the approved activities were addressing challenges proposed by program implementers who were unaware of the completed or ongoing activities. Initial linkages were made where possible to begin facilitating translation of results to programs. Missing, incomplete and misclassification of activities were some of the major challenges and a second version of the tool has been developed and is being implemented to address these challenges. The tool supports good stewardship of research and evaluation, reduces redundancy, and capacitates global health practitioners to leverage and build on existing evaluation and research efforts and bridge a gap between routine and non-routine data. While this inventory is not yet comprehensive, the utility of the inventory for characterizing relevant research and evaluation and facilitating accelerated access of results to PEPFAR programs were demonstrated.

**PN 47 A Framework for Accelerating Access to Science, Technology and Evaluation Results (FAASTER) for Global Health Impact**

## **PC 171 - Panel Contribution**

*A. Djapovic*<sup>1</sup>

<sup>1</sup> *United States Agency for International Development, Bureau for Global Health, Arlington, USA*

Ms. Djapovic Scholl will lead the panel and provide overview of the challenges and efforts within the President's Emergency Plan for AIDS Relief (PEPFAR) to centrally manage and use non-routine information. She will provide details of the framework developed to facilitate the use of evaluation and innovation, and describe the tools and systems currently deployed for improve classification of research and evaluation activities and the use of these activities alongside other types of data.

**PN 47 A Framework for Accelerating Access to Science, Technology and Evaluation Results (FAASTER) for Global Health Impact**

## **PC 172 - Panel Contribution**

*Y.F. Obeng-Aduasare<sup>1</sup>*

<sup>1</sup> *Peace Corps, Office of Global Health and HIV, Data Analytics, Washington, D.C., USA*

Ms. Obeng-Aduasare performed the data manipulation, imputation, and visualizations for the abstract submission. She will discuss considerations for secondary data use, health information systems requirements for data repositories, and HIS capacity building in resource-limited settings.

**PN 47 A Framework for Accelerating Access to Science, Technology and Evaluation Results (FAASTER) for Global Health Impact**

## **PC 173 - Panel Contribution**

*E.T. Galloway<sup>1</sup>*

<sup>1</sup> *U.S. Centers for Disease Control and Prevention, Division of Global HIV and TB, Atlanta, USA*

Ms. Galloway contributed to the quality assurance and improvement of evaluation and optimization of evaluation tracking both within her agency, U.S. CDC, and within the President's Emergency Plan for AIDS Relief (PEPFAR) interagency space. She will discuss some of the interventions deployed to enhance the quality and use of evaluations for improved policy and program planning.

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 48 Data collection in dynamic, fragile and complex environment: learning from the field**

## Data Collection in Dynamic, Fragile and Complex Environment: Learning from the Field

*A. Kalugampitiya<sup>1</sup>, R. Balinder Nanda<sup>2</sup>, R. Agrawal<sup>2</sup>*

<sup>1</sup> EvalPartners, EvalPartners, Colombo 5, Sri Lanka

<sup>2</sup> EvalGender+ and Evaluation community of India, EvalGender+ and Evaluation community of India, New Delhi, India

Predictability and certainty are pretty much relics of the past. We live and operate in a complex, dynamic environment and an increasingly inter-connected world. How do we live well with the unknown and the potential of a climate or other natural disaster in our own locales? How can we be in relationship with the unexpected – be it extreme weather, mass killings or sudden shifts in governmental policies? In these turbulent times, what can be the role of the evaluation community. Old ways are not helping to effectively deal with the changing environment. As inequities, conflicts are growing, how can evaluators provide authentic data and evidence to the policy makers and the communities which can help them in taking the right decisions? The three panelists will present their experiences on how did they do things differently in terms of collecting data, what worked, what were the challenges. We see this panel as an opportunity to not only share but also learn from the experiences of the attendees so that we can co-create knowledge.

**PN 48 Data collection in dynamic, fragile and complex environment: learning from the field**

## PC 174 - Community Based Information Collection and Support System in Afghanistan

*A. Kalugampitiya<sup>1</sup>*

<sup>1</sup> EvalPartners, EvalPartners, Colombo 5, Sri Lanka

In Afghanistan, violence against women, girls and children is very high compared to any other country. Particularly women and girls are considered as property of men and subject to violence. Women are allowed to walk out of the house with an adult man: father, husband or the elder brother. If women report any violence to police, authorities or any outsiders, their situation can become worst. Ministry of Women's Affairs (MoWA), United Nations and development organizations provide services to affected women and girls including running centers for them. However it is difficult for women and girls to access these as do not have free moment. The armed conflict deepens the situation.

In this context many organizations collect data on women and girls using community based peer groups so that information about vulnerable or affected women and girls are collected. The presenter will explain the community based peer groups in the presentation. The women's centers are located in secure and non-visible locations therefore less risk for threat from protesters. The information about women seek services from centers and affected women are kept in confidential manner. Even implementing organizations do not provide this information to the donors of the centers or the government. Ministry of Women's Affairs collect only numbers against their data needs/ indicators from implementers. However MoWA has country wide services system through Women's Affairs Officers.

**PN 48 Data collection in dynamic, fragile and complex environment: learning from the field**

## **PC 175 - Participatory Statistics to Examine the Impact of Interventions to Eradicate Modern Slavery**

*R. Balinder Nanda<sup>1</sup>*

<sup>1</sup> *EvalGender+ and Evaluation community of Indi, EvalGender+ and Evaluation community of India, New Delhi, India*

The Institute of Development Studies has been carrying out a programme of research, learning and evaluation in relation to the Freedom Fund 'hotspot' in northern India, a project that seeks to reduce bonded labour in the states of Uttar Pradesh and Bihar. The aim of the study was to estimate the prevalence of bonded labour in the selected intervention communities of the Freedom Fund hotspot in northern India. The work for this baseline study build upon scoping visits comprising interviews with NGOs, focus groups with community members, field observations, the participatory collection and analysis of 353 life stories to identify the most significant indicators of change, and the generation of a baseline of participatory statistics of 3466 households across 82 hamlets in locations covered by 14 NGOs. This was followed by a systemic action research programme which combined stakeholders analysing and developing solutions to their problems with follow-up participatory statistical analysis. We will conduct an end-line survey this year after the data collection for this study has been completed. The panelist will present the challenges of estimating changes in the magnitude of various forms of slavery; the potential of combining participatory approaches with statistical principles to generate robust data for assessing impact of slavery eradication; and the practical and ethical questions in relation to working with people living within a context of modern slavery.

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 49 Our Learnings From Ten Years' Experience Using Three Different Kinds Of Evaluative Rubrics In Our Evaluation Practice: Practical Tips**

## Our Learnings from Ten Years' Experience Using Three Different Kinds of Evaluative Rubrics in Our Evaluation Practice: Practical Tips

J. Oakden<sup>1</sup>, N. Wehipeihana<sup>1</sup>, J. King<sup>2</sup>, K. McKegg<sup>3</sup>

<sup>1</sup> Kinnect Group, Director, Wellington, New Zealand

<sup>2</sup> Kinnect Group, Director, Auckland, New Zealand

<sup>3</sup> Kinnect Group, Director, Hamilton, New Zealand

**Rationale:** Evaluation rubrics are a powerful and influential approach to evaluation-specific methodology that can be used in collaborative/participatory or independent evaluations to build a clear, shared understanding of how quality, value, and effectiveness are defined. Evaluative rubrics make explicit the basis for evaluative judgments about effectiveness or performance, as well as importance.

Drawing from their experience of using rubrics in many evaluation settings, the presenters in this panel session will provide an overview of rubrics, as well as more detail about different kind of rubrics and their uses, their strengths and weaknesses, and the ability of rubrics to explore and integrate shared values providing a clear and transparent basis for making decisions.

**Objectives:** In this panel presentation, participants will gain insights from panel members' practice about rubrics in the following areas:

An overview of rubrics

- What are rubrics?
- Where do they come from?
- What are the components of a rubric?
- Why are they useful / transformative for evaluation practice?

Different kinds of rubrics:

- What different types of rubrics are there?
- What are their key features?
- What are the design considerations for each?
- What is the comparative value of each type for making evaluative judgments?
- The strengths and weaknesses of rubrics
- What are the strengths of rubrics?
- Troubleshooting, faults and mishaps – overcoming the weaknesses of rubrics in practice?
- How do they transform evaluation practice?

Using rubrics to integrate shared values

- Whose perspectives and values count when using rubrics?
- How do you weave different values into the design and use of a rubric?
- Why does this matter?

**Relevance to the conference strands and theme:** Evaluative rubrics help evaluations generate valid, credible insights that decision makers and other stakeholders can trust and use to take action. They are now being widely used across diverse and complex settings to engage stakeholders in deep values and evidence-based conversations about quality, value and importance as well as to support participatory synthesis and evaluative interpretation of qualitative and quantitative evidence.

They are being used not just for the evaluation of individual programs, but also for developing a shared language and interpretive frame across multiple programs and policy initiatives in local, regional, national and global settings. From the smallest and simplest evaluation, to the most complex and uncertain situations they can be incredibly powerful for building a shared understanding about what is valued, about "how good is good"; sometimes even challenging traditional power relations. This panel will discuss the diverse uses of rubrics that span a wide range of contexts, stakeholders, and applications. This includes for example the use of rubrics in international development, environmental, education and health settings as well as indigenous, legislative and policy settings.

**PN 49 Our Learnings From Ten Years' Experience Using Three Different Kinds Of Evaluative Rubrics In Our Evaluation Practice: Practical Tips**

## **PC 180 - An Overview of What Evaluative Rubrics Are**

*N. Wehipaihana*<sup>1</sup>

<sup>1</sup> Kinnect Group, Director, Wellington, New Zealand

Getting Oriented – An Overview Of What Evaluative Rubrics Are This presentation will cover: What are rubrics? We will provide a brief overview of the theory behind rubrics. Where do they come from? We will explore their origins. What are the components of a rubric? This section will introduce the three components of rubrics The final section will cover why we believe evaluative rubrics are useful / transformative for evaluation practice?

**PN 49 Our Learnings From Ten Years' Experience Using Three Different Kinds Of Evaluative Rubrics In Our Evaluation Practice: Practical Tips**

## **PC 181 - Exploring Three Different Kinds of Evaluative Rubrics**

*J. Oakden*<sup>1</sup>, *J. King*<sup>1</sup>

<sup>1</sup> Kinnect Group, Director, Wellington, New Zealand

Exploring Three Different Kinds of Evaluative Rubrics The presentation will cover:

What different types of rubrics are there? Three types will be discussed, analytic, generic and holistic rubrics.

What are their key features? We will look at the components that make up each kind of rubric.

What are the design considerations for each? We will explore what it takes to develop and use rubrics through the evaluation process.

What is the comparative value of each type for making evaluative judgments? Some rubrics provide more certainty, others more flexibility. We will explore these differences.

**PN 49 Our Learnings From Ten Years' Experience Using Three Different Kinds Of Evaluative Rubrics  
In Our Evaluation Practice: Practical Tips**

## **PC 182 - Reflecting on the Strengths and Weaknesses of the Different Kinds of Evaluative Rubrics**

J. King<sup>1</sup>

<sup>1</sup> Kinnect Group, Director, Auckland, New Zealand

Reflecting on the strengths and weaknesses of the different kinds of evaluative rubrics: What are the strengths of rubrics? The presenter will discuss some of the benefits we have found in our evaluation practice using rubrics. But they are not always easy to do well. Therefore we include a section on troubleshooting, faults and mishaps – to overcome the weaknesses of rubrics in practice. This part of the presentation will cover the challenges we have found using rubrics, and provide some pointers to help you on your way. How do evaluative rubrics transform evaluation practice? We consider evaluative rubrics our evaluative equivalent of GPS for navigating our way through an evaluation. We will show you why.

**PN 49 Our Learnings From Ten Years' Experience Using Three Different Kinds Of Evaluative Rubrics  
In Our Evaluation Practice: Practical Tips**

## **PC 183 - How We Might Use Evaluative Rubrics to Integrate Shared Values into Evaluation**

K. McKegg<sup>1</sup>

<sup>1</sup> Kinnect Group, Director, Hamilton, New Zealand

How we might use evaluative rubrics to integrate shared values into evaluation? This presentation explores: Whose perspectives and values count when using rubrics? We have found that stakeholders often have differing views, and exploring and navigating these differences can be important throughout the evaluation. This leads us to explore how we might weave different values into the design and use of an evaluative rubric? When a range of values are included, it makes the evaluation more relevant and valuable to a broader range of users. Why this matters? The use of evaluative rubrics is a way to systematically embed the hopes and aspirations of a range of service users and stakeholders, including those who may not often have a voice in the evaluation.

Thursday, 4 October 2018  
11:30–13:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 50 Using Outcome Harvesting to Learn about what Supports Resilience**

## Using Outcome Harvesting to Learn about what Supports Resilience

G. Scheers<sup>1</sup>, R. Wilson-Grau<sup>2</sup>, R. Smith<sup>3</sup>, M. Smith<sup>4</sup>, G. Blundo-Canto<sup>5</sup>

<sup>1</sup> Goele Scheers Consultancy, Independent Consultant, Gentbrugge, Belgium

<sup>2</sup> Ricardo Wilson-Grau Consultoria em Gestão Empresarial Ltda, Independent Consultant, Tijuca, Brazil

<sup>3</sup> RDS consulting, Independent Consultant, London, United Kingdom

<sup>4</sup> Independent Consultant, Independent Consultant, Stockholm, Sweden

<sup>5</sup> French Agricultural Research Centre for International Development CIRAD, Researcher, Montpellier, France

We will demonstrate why in the last six years Outcome Harvesting has attracted mainstream interest as an evaluation approach for development interventions that foment resilience without fundamental loss of identity. The Ford Foundation, UNDP, USAID and the World Bank have or are piloting and recommending the approach for learning about achievements of interventions operating in dynamic, uncertain and unpredictable situations. After an introduction by Ricardo Wilson-Grau, whose book *Outcome Harvesting for Monitoring and Evaluation: – Practical Applications of Essential Principles* will be published by IAP to coincide with the EES 2018 Conference, four colleagues will present their experiences. Goele Scheers, an independent consultant based in Belgium, supported Free Press Unlimited (FPU) to use Outcome Harvesting for evaluations in Somalia and Pakistan. FPU helps local journalists and media professionals in war zones and conflict areas to provide their audience with trustworthy news and information, through trainings, emergency support and capacity building. In 2017, FPU piloted Outcome Harvesting applying the six steps differently each time in order to draw lessons on which approach would best fit the needs of the organisation. In Pakistan, for example, the outcomes were harvested via e-mail; in Somalia, the outcomes were harvested in a workshop. Genowefa Blundo Canto, a researcher in impact assessment with the French Agricultural Research Centre for International Development (CIRAD), adapted the Outcome Harvesting tool for the systematic evaluation of multiple outcomes of agricultural research for development, based on her experience at the International Centre for Tropical Agriculture (CIAT). Beginning In 2014, CIAT customised the Outcome Harvesting tool for yearly outcome evaluations. As a result, CIAT enhanced its adaptive management through better understanding of how the institution interacts with actors who use its agricultural research to improve agricultural systems, identifying successful strategies and those that should be improved or changed. Mariam Smith, an independent consultant based in Sweden, engaged in an Outcome Harvesting evaluation that supported indigenous programme staff of a human rights intervention in Cambodia. The evaluation enabled the staff to learn from, and communicate about, change they were influencing in support of indigenous people demanding their rights in the wider society. Unlike some previous evaluations, staff and indigenous people alike gained a more vivid picture of both positive and negative outcomes, expected and unexpected, without feeling threatened and losing dignity. In addition, the outcome focused approach was strengthened among all staff. Richard Smith, an independent consultant based in the UK, applied Outcome Harvesting to evaluate the Global Network of Civil Society Organisations for Disaster Reduction (GNDR). A variety of Outcome Harvesting processes were used – including document review, surveys, focus groups with network members and engaging network secretariat informants – with the innovation of first seeking 'basic perceived outcomes' that were, where possible, then developed into robust, SMART outcome descriptions. The evaluation a) described outcomes for the first time, b) used outcomes achieved to assess progress towards pre-determined objectives, and c) provided a participatory foundation for strategy and monitoring system development.

**PN 50 Using Outcome Harvesting to Learn about what Supports Resilience****PC 184 - Lessons Learned from Piloting Outcome Harvesting for Free Press Unlimited (FPU) in Somalia and Pakistan**

G. Scheers<sup>1</sup>

<sup>1</sup> Goele Scheers Consultancy, Independent Consultant, Gentbrugge, Belgium

Free Press Unlimited (FPU) helps local journalists in war zones and conflict areas to provide their audience with trustworthy news and information. In almost 40 countries Free Press Unlimited supports journalists and media professionals with trainings, emergency support and capacity building. Starting in 2017, FPU piloted Outcome Harvesting in 2 of these countries with their partners. In each country, the application of the six Outcome Harvesting steps was done differently to be able to draw lessons on which approach would best fit the needs of the organisation. In Pakistan for example, the outcomes were harvested via e-mail, whereas in the case of Somalia we harvested the outcomes in a workshop. This presentation will highlight the main lessons learned from this pilot.

**PN 50 Using Outcome Harvesting to Learn about what Supports Resilience****PC 185 - Using Outcome Harvesting to Describe Results, Support Learning and Evaluate a Civil Society Network for Disaster Reduction**

R. Smith<sup>1</sup>

<sup>1</sup> RDS consulting, Independent Consultant, London, United Kingdom

The Global Network of Civil Society Organisations for Disaster Reduction (GNDR) aims to help its members improve the lives of people affected by disasters world-wide. During the period evaluated, the network had sought a mix of internal and external results: i) further strengthen relationships within the network; ii) enhance engagement between civil society and other key stakeholders; and iii) develop a programme to communicate the realities of people most at risk from disaster. A variety of processes were used to detect possible outcomes from across the network, including document review, surveys, engaging internal informants and focus groups with network members. Working with internal informants, 'outcome signals' were, where possible, developed into robust, SMART outcome descriptions. The harvest was used to

- a) describe outcome-level results for the first time,
- b) support learning, participation and ownership across the network and
- c) evaluate planned and unintended outcomes in relation to pre-determined objectives.

**PN 50 Using Outcome Harvesting to Learn about what Supports Resilience****PC 186 - Adapting the Outcome Harvesting tool for the systematic evaluation of multiple outcomes of agricultural research for development**

G. Blundo-Canto<sup>1</sup>

<sup>1</sup> French Agricultural Research Centre for International Development CIRAD, Researcher in Impact Assessment, Montpellier, France

Agricultural research for development organizations are under pressure to provide evidence of developmental outcomes in light of global environmental and developmental concerns. In 2014, the International Centre for Tropical Development (CIAT) developed a monitoring, evaluation and learning (MEL) system that applies aspects of the Outcome Harvesting tool for yearly outcome evaluations. The system focuses on learning for adaptation and improvement of innovation processes. Three evaluations have been carried out since 2014, and a large-scale

evaluation adapting Outcome Harvesting will be carried out in 2018. We discuss the ability of the MEL system that uses a customization of Outcome Harvesting to provide insight into the processes and interactions with actors that generate outcomes. Lessons learned from the evaluations have allowed adaptive management and better understanding of how the research institution interacts with actors, the strategies that appeared to be successful and those that should be improved or changed.

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 51 Evaluation Criteria for the 21<sup>st</sup> Century: Panel 1 Rethinking Our Evaluation Criteria**

## **Evaluation Criteria for the 21<sup>st</sup> Century: Panel 1 Rethinking Our Evaluation Criteria**

*J. Puri*<sup>1</sup>

<sup>1</sup> *Green Climate Fund, Independent Evaluation Unit, Songdo, Republic of Korea*

This special symposium consists of two consecutive, interlinked panels. They are based on the assumption that it is essential that the global evaluation community fundamentally reconsiders the evaluation criteria that play such an important role in shaping our work. Critical interrogation of our current practices in the design and application of evaluation criteria is essential and urgent to ensure that we do not submit to convenient arguments about their framing, and to comfort zones in their application. They have to be appropriate for the era in which we now have to conduct our work. Evaluation criteria are of critical importance to the evaluation enterprise. They are those qualities or dimensions on which evaluators judge the extent to which something is successful, good, effective, useful, a good return on investment, and so on. They provide the yardsticks through which performance or accomplishments are viewed and measured. Evaluation criteria are context-sensitive and value based, reflecting the normative beliefs that stakeholders, shareholders and/or evaluators in a particular situation hold about that which is being assessed. For every evaluation the evaluator has the professional responsibility to use his/her expertise to determine and use the criteria appropriately for the circumstances. We live in an era defined by the intersection of the Sustainable Development Goals, the Fourth Industrial Revolution, calls for more a humane and planet-centric global economy, and many macro influences sweeping the planet, from climate change to increasing dissatisfaction with existing economic and political systems to increasingly severe geopolitical competition for power and resources. Given the importance of evaluation criteria, what should they look like to ensure that our evaluations respond to the opportunities and challenges that the world faces today? On what basis should we select and apply evaluation criteria? How can we ensure their appropriate use so that they can best serve the lives, societies and ecosystems we are trying to improve? How do we balance pragmatic considerations with ideals about evaluation that works for development? This is of particular importance when considering the ubiquitous 'DAC criteria' – relevance, efficiency, effectiveness, impact and sustainability – which for more than two decades have been widely used, especially in the Global South. They are widely recognized as having been very useful in many ways, yet tend to determine to a large extent the evaluation questions, and hence the focus of the evaluation. Some have questioned the desirability of having such a powerful force directing the focus and therefore the findings of many evaluations around the world. It has become important to review their value and utility for evaluation in this era. We have brought together two panels to discuss and debate key issues around evaluation criteria in general and the DAC criteria in particular. We believe that this is part of considering Doing Evaluation Differently. We all need to take time to examine and, if necessary, reform or transform our evaluation practice and system to contribute effectively to resilient societies and ecosystems in the 21<sup>st</sup> Century.

### PANEL 1. RETHINKING OUR EVALUATION CRITERIA

From very different perspectives and experiences, this panel will discuss the conceptualization and use of evaluation criteria that are suitable for the era in which we now live – based on our understanding of what needs to be evaluated, for what purpose, and informed by which frameworks, principles or values.

Panel members will among others consider

- (i) the normative nature of sustainable development, with particular reference to the Global South;
- (ii) the need for transformative change at national and global level;
- (iii) development viewed through a complex systems lens; and

(iv) private sector engagement with social good and social impact measurement.

The presentations, followed by comments and debates from the floor, will provide food for thought and set the scene for the discussion of the DAC criteria by Panel 2.

**PN 51 Evaluation Criteria for the 21<sup>st</sup> Century: Panel 1 Rethinking Our Evaluation Criteria**

**PC 187 - Criteria, Values and the Logic of Evaluation**

S. Sridharan<sup>1</sup>

<sup>1</sup> University of Toronto, Department of Health Policy- Management and Evaluation, Toronto, Canada

Sanjeev Sridarhan's contribution How each panelist will contribute and for how long: Each panelist will give a presentation of 10 minutes, followed by an open discussion and debate that will be continued as part of the discussions after Panel 2.

**PN 51 Evaluation Criteria for the 21<sup>st</sup> Century: Panel 1 Rethinking Our Evaluation Criteria**

**PC 188 - Evaluation Criteria: Perspectives from the Global South**

S. Zaveri<sup>1</sup>

<sup>1</sup> Community of Evaluators South Asia, Founder member and Vice President, New Delhi, India

Sonal Zaveri's contribution.

**PN 51 Evaluation Criteria for the 21<sup>st</sup> Century: Panel 1 Rethinking Our Evaluation Criteria**

**PC 189 - Criteria for the Evaluation of Sustainable Development**

Z. Ofir<sup>1</sup>

<sup>1</sup> Stellenbosch University, Honorary Professor, Stellenbosch, South Africa

Zenda Ofir's contribution.

**PN 51 Evaluation Criteria for the 21<sup>st</sup> Century: Panel 1 Rethinking Our Evaluation Criteria**

**PC 190 - Criteria on the Interface of Private Sector and Development Evaluation**

A. Lord<sup>1</sup>

<sup>1</sup> Permian Global, monitoring- evaluation and learning programme, London, United Kingdom

Alex Lord's contribution.

**PN 51 Evaluation Criteria for the 21<sup>st</sup> Century: Panel 1 Rethinking Our Evaluation Criteria**

**PC 191 - Discussant: Challenges for the Evaluation Profession in Redefining Criteria**

R.D. van den Berg<sup>1</sup>

<sup>1</sup> IDEAS, President, Leidschendam, the Netherlands

Rob D. van den Berg's contribution.

Thursday, 4 October 2018  
17:30–19:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 52 Evaluation in Turbulent Times. Pushing Boundaries through a Complexity Lens**

## Evaluation in Turbulent Times. Pushing Boundaries Through a Complexity Lens

M. Arranz<sup>1</sup>, M. Bamberger<sup>2</sup>, S. Chaplowe<sup>3</sup>, E. Raimondo<sup>4</sup>

<sup>1</sup> Oxfam Great Britain, Planning- Monitoring- Evaluation- Learning, Oxford, United Kingdom

<sup>2</sup> Independent Consultant, Independent, Beaverton, USA

<sup>3</sup> Children's Investment Fund Foundation, Evidence- Monitoring and Evaluation Climate, London, United Kingdom

<sup>4</sup> World Bank, Independent Evaluation Group, Washington DC, USA

Reality is messy and mostly unpredictable. Political, economic and social changes are complex phenomena that do not follow linear causal pathways. We have learned from complexity theory that in complex situations 'change follows an unpredictable trajectory... Cause and effect is only coherent retrospectively and cannot be repeated' (Guijt, 2010: 287).

Interventions aiming at social and policy change deal with four dimensions of complexity: organisational dynamics; contextual factors (knowledge about the environment in which the intervention works); nature of the evaluand; and causal complexity (difficulty of understanding the mechanisms that may connect causes to desired effects).

Complexity theory and systems thinking are increasingly permeating development programmes and evaluation has the potential to help or hinder this trend.

In a globalised world, systems are more than ever interconnected and shocks generate ripple effects that are difficult to anticipate. In turbulent times and volatile environments, evaluation more than ever needs to provide a guiding light to understand how we can adapt to these shocks and be more resilient to them. This might require a high degree of bravery and curiosity, to push pre-established boundaries and to look beyond the intended effects.

Complexity-aware monitoring and evaluation (M&E) invites us to explore how the context is affecting the intervention and the how different elements in the system are interrelated. Interest is paid to unintended effects as much as to intended ones, and the assumptions that underpin the Theory of Change are tested.

Results-based and performance-driven M&E systems can ignore contextual and casual complexity and simplify causal relationships. These approaches concentrate the attention mostly in the fidelity of the design and the achievement of expected results, failing to offer a more holistic picture of what changes are really occurring, for whom, and why. And yet, they are still widely preferred by many organisations, donors and implementers. As evaluators seeking to provide a "guiding light," we must also be careful not to over-simplify the multiplicity of the interdependent factors and nonlinear change that characterise the operational environments services are delivered (and evaluated).

This panel will present some successful experiences introducing a complexity lens to M&E and will discuss how brave evaluators can broaden the scope and utility of evaluations by pushing their boundaries. It will also explore how to overcome frequent organisational, political and methodological factors that generate resistance to complexity-aware M&E.

The chair will frame the session, followed by two presentations that will offer complementary angles and by an additional discussant. The first panellist will present a framework to measure indirect benefits and will share learning about how organisational-level metrics can incentivise complexity-aware MEL systems for programmes. The second panellist will examine the practicalities of applying a complexity lens during a real-time evaluation of a humanitarian operation responding to the European migration crisis in 2016. The additional discussant will provide an institutional perspective of the challenges of introducing a complexity lens in independent evaluation.

**PN 52 Evaluation in Turbulent Times. Pushing Boundaries through a Complexity Lens****PC 192 - Who Is Benefiting? An Approach to Measure Indirect Effects**

*M. Arranz*<sup>1</sup>

<sup>1</sup> *Oxfam Great Britain, Planning- Monitoring- Evaluation- Learning, Oxford, United Kingdom*

This presentation will explore how a M&E framework initially intended at estimating indirect reach became an essential to understanding the programme's Theory of Change, its scale up pathways and indirect effects. The theory-based conceptual and methodological framework to estimating who is benefiting indirectly from the programme will be presented, illustrating the steps and providing some practical examples of programmes that have used it. The presentation will also explore how this approach is supporting the incorporation of a complexity-lens to monitoring and evaluation, especially for influencing work, and how this framework is being used to inform design as much as adaptation during the implementation.

**PN 52 Evaluation in Turbulent Times. Pushing Boundaries through a Complexity Lens****PC 193 - Navigating Complexity in the Real Time Evaluation of Humanitarian Response to European Migration**

*S. Chaplowe*<sup>1</sup>

<sup>1</sup> *Children's Investment Fund Foundation, Evidence- Monitoring and Evaluation Climate, London, United Kingdom*

Since 2010, the IFRC adopted Real-Time Evaluations (RTE) to assess its emergency operations. In the summer of 2016, the IFRC commissioned a team to assess the overall effectiveness of its 2015–2016 European Migration Response. The concepts of complexity and systems thinking has had a growing influence on the understanding and practice of evaluation (Bamberger et al., 2015; Hargreaves, 2010; Morell, 2010; Patton, 2011; Ramalingam et al., 2008; and Williams & Hummelbrunner, 2009). This RTE highlighted the relevance of this discourse because the European migration context is complex, with a multiplicity of interdependent factors and nonlinear change, resulting in considerable ambiguity and uncertainty. This required the RTE team to adopt a systems perspective to best encompass the diversity of interconnected factors and actors affecting the European migration context and response. Migration does not exist in a political vacuum, and assessment of the response needed to analyse the configuration of economic, political, and social factors at the regional as well as national levels. This included divergent and changing beliefs and opinions in host countries, especially as migration became politicised and xenophobia intensified by vocal nationalist and anti-immigrant parties, and accentuated by economic hardship following the 2008 recession. This presentation will draw upon lessons of this complexity-aware Real-Time Evaluation, also published as an article in the October 2017 EES Evaluation Connections.

Thursday, 4 October 2018  
17:30–19:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 53 Making Sense of Outcomes: Using the signification process of SenseMaker to enrich Outcome Mapping and Outcome Harvesting**

## **Making Sense of Outcomes: Using the Signification Process of SenseMaker to Enrich Outcome Mapping and Outcome Harvesting**

*R. Wilson-Grau*<sup>1</sup>

<sup>1</sup> *Ricardo Wilson-Grau Consultoria em Gestão Empresarial Ltda, Consultancy, Rio de Janeiro, Brazil*

This panel session will illustrate how a signification process – as it is commonly applied in the SenseMaker practice – can be used to add new layers of meaning through visual pattern analysis of outcomes that are generated through an Outcome Mapping or Outcome Harvesting process. The particular way of analysis facilitates richer analysis in the monitoring and evaluation process. Outcome Mapping is an actor-centered PM&E approach that focuses on the behavioral change of those actors / partners that are directly influenced by the programme intervention team, i.e. the so called boundary partners. During the design stage, intentional outcomes are identified for each boundary partner and formulated as outcome challenges and progress markers. During the monitoring process, the programme team and boundary partners keep track of emerging changes in behaviour – through systematic data collection and reflection – in the lifespan of the programme. Outcome Harvesting is a utilization-focused, participatory tool that enables evaluators to identify, formulate, verify, and make sense of outcomes they have influenced when relationships of cause and effect are unknown. Unlike some evaluation methods, Outcome Harvesting does not measure progress towards predetermined outcomes or objectives, but rather collects evidence of what has been achieved, and works backward to determine whether and how the project or intervention contributed to the change. SenseMaker is an innovative narrative-based research and monitoring tool that helps decision-makers to make sense of complex processes and systems in order to act in it. It is developed by Cognitive Edge and based on mass capture of fragmented material, mostly micro-narratives that are self-signified by the storytellers. Beneficiaries themselves then assign meaning to their own stories using a set of predefined “signifiers”, rather than an external intermediary assigning meaning to their stories, as is the case in conventional qualitative analysis. SenseMaker, combines the advantage of self-signified stories with the value of doing large-scale statistical analysis that allow the exploration of patterns. Understanding individual outcomes are interesting, but the real power for understanding reality lays in an analysis of the entire set of outcomes and the detection of patterns of change over time to generate insights on progress and further action planning. The use of the signification of SenseMaker during the harvesting or (collective) analysis enhances pattern visualisation and reveal insights that go beyond what can be read in the outcome fragments or statements. A first part will provide an introduction in the conceptual underpinning of the signification approach, and how it can be used in the design, collection and analysis steps of the OM or OH monitoring and evaluation process. The second part will showcase how a ActionAid programme used signifier questions to enrich the analysis.

**PN 53 Making Sense of Outcomes: Using the signification process of SenseMaker to enrich Outcome Mapping and Outcome Harvesting**

**PC 194 - Signification of Outcomes: Enhancing The Harvesting and Analysis of Outcomes**

S. Deprez<sup>1</sup>

<sup>1</sup> *Voices That Count, Training & Consultancy, Kessel-Lo Leuven, Belgium*

The first session will provide an introduction in the conceptual underpinning of the signification approach and how generic and tailor-made signifier questions can be used in the design and collective harvesting of outcomes in an OH/OM process. It will also highlight how it enhances the analysis step including participatory analysis and collective interpretation workshops.

**PN 53 Making Sense of Outcomes: Using the signification process of SenseMaker to enrich Outcome Mapping and Outcome Harvesting**

**PC 195 - Using Signifier Questions in the Analysis of the Outcome Harvesting Evaluation of Actionaid Denmark's Tax Justice Programme**

G. Scheers<sup>1</sup>, H. Bach<sup>2</sup>

<sup>1</sup> *Goele Scheers Consulting, Training & Consulting, Gent, Belgium*

<sup>2</sup> *ActionAid, Copenhagen, Denmark*

In this session we will showcase how signifier questions enriched the analysis in the Outcome Harvesting evaluation of ActionAid Denmark's Tax Justice Programme. We will demonstrate how the signifier questions were entered into the system and how this data was used to answer, and give more in depth meaning to, the evaluation questions. Furthermore, we will discuss the advantages and challenges of this approach that we experienced during this process.

Thursday, 4 October 2018  
17:30–19:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

#### PN 54 Understanding How Adaptive Management Can Support Resilience Strengthening

### Understanding How Adaptive Management Can Support Resilience Strengthening

L. Wingender<sup>1</sup>

<sup>1</sup> Senior Advisor, Program Performance and Quality, Mercy Corps, USA

A community's or individual's ability to be resilient varies depending on the type and magnitude of a threat (the exact nature of which cannot be known in advance) and this presents a challenge during program design and implementation, particularly when it comes to evaluation. Working to strengthen resilience requires teams to acknowledge uncertainty and unpredictability, seek opportunities to experiment and understand what works, and adapt and scale approaches in response to rapid feedback loops, changing circumstances and new information. Furthermore, an approaching threat requires extraordinary preparation, and flexibility to adapt program activities and indicators. The role of monitoring and evaluation in these circumstances is the focus of this panel.

Resilience measurement is a relatively young field and one in which resilience practitioners – both evaluators and decision-makers – are investing in rapidly learning to respond and measure effectively when a shock or stress strikes. In late 2017, The Resilience Measurement, Evidence and Learning Community of Practice launched four innovation awards for initiatives that promise to move the resilience measurement, evidence, and practice field forward. One of the innovation awards focuses on the challenge faced by program managers of resilience programming: in the complexity of resilience programming, they have to make reliable and quick decisions to strengthen or maintain resilience in the face of shocks and stresses, but data and analyses from robust evaluation exercises may not be timely or able to provide the information needed to make decisions.

A team comprised of Itad and Mercy Corps lead a scoping study that explores the decision-making process of resilience program managers, their information needs to adaptively manage their programs, what information they receive and how they interpret and use it. Focusing on the user can provide feedback to monitoring and evaluation experts to design simpler yet robust tools that decision-makers and their monitoring and evaluation staff can use throughout a project life cycle, without having to wait for baseline or endline evaluation results.

This panel will explore the needs of decision-makers as interpreters and users of information; showcase two key examples of how development agencies are attempting to design simpler resilience measurement tools for adaptive management; and provide recommendations to evaluators on how the M&E field could change to meet the needs of decision makers. The examples will focus on both design and implementation and how programs are using resilience information to adapt program activities and will draw on programs' experience in incorporating resilience measurement thinking in the design of results chain and data review processes.

#### PN 54 Understanding How Adaptive Management Can Support Resilience Strengthening

### PC 196 - Panel Contribution

B. Murphy<sup>1</sup>

<sup>1</sup> ITAD, Brighton, United Kingdom

Ben is leading a scoping study exploring the linkages between adaptive management and resilience measurement. He will talk about the findings emerging from this study.

**PN 54 Understanding How Adaptive Management Can Support Resilience Strengthening**

**PC 198 - Panel Contribution**

B. Sladkova<sup>1</sup>

<sup>1</sup> *Itad, Consultant, Brighton, UK*

Bara is a researcher on a scoping study exploring the linkages between adaptive management and resilience measurement. She will talk about the findings emerging from this study.

**PN 54 Understanding How Adaptive Management Can Support Resilience Strengthening**

**PC 199 - Panel Contribution**

J. Scantlan<sup>1</sup>

<sup>1</sup> *Mercy Corps, Technical Support Unit - Resilience, Portland, USA*

Jill will contribute to the panel by sharing her experience developing resilience measurement systems in Nepal, Mongolia and Timor-Leste. She will focus on sharing best practices for using data from measurement systems to inform decision-making within the program.

Thursday, 4 October 2018  
17:30–19:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 55 Issues and Challenges in the Establishment and Development of Voluntary Organisations for Professional Evaluation (VOPE)**

## Issues and Challenges in the Establishment and Development of Voluntary Organisations for Professional Evaluation (VOPE)

*B. Cousins<sup>1</sup>*

<sup>1</sup> CRECS University of Ottawa, Faculty of Education, Ottawa- ON, Canada

**Rationale:** National and regional voluntary organizations for professional evaluation (VOPEs) have considerable potential to enhance evaluation capacity development in the interest of leveraging desirable social change. Such organizations play a significant role in legitimizing and enhancing the credibility of evaluation; increasing evaluation use; building evaluation capacity; fostering the acceptance of guiding principles and codes of ethics; providing fora for exchange on best practices; and increasing support for cultural specificity with regard to evaluation design and practice (Quesnel, 2006). Some would argue that the sustainability of these organizations is therefore critical as evaluation's role in navigating critical times increases. Yet, despite the availability of some practical wisdom and resources, little is known about the acute challenges facing newly established and developing VOPEs and about potential solutions that may contribute to their viability and sustainability.

**Objectives:** The panel seeks to identify and elucidate

- (i) challenges facing newly established and developing VOPEs, and
- (ii) strategies and solutions that are likely to be viable in meeting such challenges and fostering VOPEs' viability and sustainability.

**Narrative/Justification:** Globally, the establishment and development of national and regional VOPEs have increased dramatically (Segone, 2006). While some major national and regional evaluation associations and societies have been in operation for decades, these are far outnumbered by new and developing evaluation organisations. While the research base on evaluation capacity building (ECB) has developed significantly over the past decades, scholarship on new and emerging VOPEs remains limited. Yet, through the hard work of multilateral partnership organizations like EvalPartners and the International Organization for Cooperation and Evaluation (IOCE), a range of practical resources has been made available. Edited volumes such as those published by Segone and Ocampo (2006) and Rugh and Segone (nd-2013) have provided case reports of the development experiences of VOPEs ranging from those which are institutionalized to those in the early stages of development. However, such publications only indirectly engage with the issue of challenges facing newly established VOPEs and potentially viable solutions to those challenges. This panel will engage with these issues more directly. The panel will build upon the experiences of two recently established VOPEs, one at a very early stage of development, and the second at a more mature stage of institutionalization. Specifically, the experiences of the Turkish Monitoring and Evaluation Society (TMES), established in 2013, and those of the Jordanian Development Evaluation Association (EvalJordan), a VOPE established in 2014, will be shared. While some consideration will be given to the initial impetus for founding these associations, the primary focus will be on challenges facing the organizations in the early stages of operation and on various solutions that either have been demonstrated to help or have great potential to help develop the sustainability of the organizations. These presentations will be augmented by a close examination of the body of scholarship, including research and theory related to the establishment and development of VOPEs. The panel seeks to engage with audience members who may wish to share their own experiences associated with the early stages of VOPE development.

**PN 55 Issues and Challenges in the Establishment and Development of Voluntary Organisations for Professional Evaluation (VOPE)****PC 200 - Challenges and Strategies of the Turkish Monitoring and Evaluation Society**

*S. Akgüngör<sup>1</sup>, E. Volkan<sup>2</sup>*

<sup>1</sup> Dokuz Eylül University, Economics, İzmir, Turkey

<sup>2</sup> Ministry of Development of Turkey, Monitoring- Evaluation and Analysis Department, Ankara, Turkey

TEMS was established in 2013. Since then TMES aspires to develop a network of evaluators in Turkey as well as share knowledge on evaluation. TEMS aims to enhance accountability, organizational learning and development in Turkey and intends to foster a monitoring and evaluation (M&E) culture across a wide range of institutions. In December 2017, TMES held a workshop with participation of 27 invited evaluation professionals and stakeholders with a particular focus on evaluation needs and capacities in Turkey. TMES is currently conducting a survey of information needs in order to help shape decision making around its roles and responsibilities. While there exists a great deal of enthusiasm for the development of TMES, the Society's leadership is faced with a number of challenges in the early going. The major problem is that there is a lack of awareness and use of evaluation results both in public and non-government sector. Clear priority exists in developing strategies for evaluation capacity building in Turkey. The panelists will describe the country context and impetus for establishing the Society and identify the principal challenges it faces. The discussion will include the immediate, short- and long-term strategies adopted by TMES in its quest to move toward institutionalization.

**PN 55 Issues and Challenges in the Establishment and Development of Voluntary Organisations for Professional Evaluation (VOPE)****PC 201 - Challenges and Strategies of the Jordanian Development Evaluation Association**

*M. Qaryouti<sup>1</sup>*

<sup>1</sup> EvalJordan, Amman, Jordan

Legally recognized by the Jordanian government, EvalJordan was officially registered in May 2014 as voluntary society under the umbrella of the Ministry of Social Development. Although the early efforts of establishing EvalJordan focused on providing a platform for knowledge sharing and exchanging experiences the association now seeks to be an integral part of an effective decision making culture that is evidence-based, measurable and accountable, leading to successful and sustainable national development in Jordan. In its early years of development EvalJordan focused on

- (i) establishing itself and raising awareness about evaluation throughout the country
- (ii) bringing together M&E practitioners and professionals to share and exchange knowledge, expertise, lesson learned and best practices, and
- (iii) initiating capacity building opportunities for members of the network and stakeholders, especially government staff.

Having existed for four years, EvalJordan is well positioned to reflect on its early growth and development. The panelist will describe a range of challenges that the Association faced in these early years, how these challenges were addressed, and goals and priorities moving forward.

**PN 55 Issues and Challenges in the Establishment and Development of Voluntary Organisations for Professional Evaluation (VOPE)**

**PC 202 - Scholarship and Practical Resources for Voluntary Organisation for Professional Evaluation (VOPE) Development**

*H. Al Hudib<sup>1</sup>*

<sup>1</sup> CRECS University of Ottawa, Faculty of Education, Ottawa- ON, Canada

The panelist will discuss the results of a systematic review and integration of the literature conducted by Al Hudib and Cousins on what is known about VOPE establishment and development and what the ECB literature may have to offer in terms of guidance to such organizations in their early stages of development. Al Hudib's contribution will focus on scholarship and practical resources that have been made available to developing VOPEs. The objective of this review is to identify evidence-based strategies and actions that seem viable in helping VOPEs to solidify their establishment and assist in their institutionalization.

**PN 55 Issues and Challenges in the Establishment and Development of Voluntary Organisations for Professional Evaluation (VOPE)**

**PC 203 - Research and Scholarship on Organisational Evaluation Capacity Building: Insights for Developing Voluntary Organizations for Professional Evaluation (VOPEs)**

*B. Cousins<sup>1</sup>*

<sup>1</sup> CRECS University of Ottawa, Faculty of Education, Ottawa- ON, Canada

The panelist will discuss the results of a systematic review and integration of the literature conducted by Al Hudib and Cousins on what is known about VOPE establishment and development and what the ECB literature may have to offer in terms of guidance to such organizations in their early stages of development. Cousins' contribution will focus on research and scholarship on organizational ECB, specifically organizational capacity to do and use evaluation. The objective of the presentation will be to identify factors and evidence-based strategies emerging from the research that may be extrapolated to the question of VOPE establishment and development. The panelist will also offer some discussant remarks and observations arising from the foregoing presentations and will lead an exchange between audience members and the panelists.

Thursday, 4 October 2018  
17:30–19:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 56 EvalAgenda 2020 In the Region Eastern Europe, Central Asia & South Caucasus: Opportunities and Challenges for Young and Emerging Evaluators**

## EvalAgenda 2020 in the Region Eastern Europe, Central Asia & South Caucasus: Opportunities and Challenges for Young and Emerging Evaluators

*A. Rishko-Porcescu<sup>1</sup>, B. Zheenbekov<sup>2</sup>, A. Harutyunyan<sup>3</sup>, T. Narchynska<sup>4</sup>, K. Rasulova<sup>5</sup>, M. Segone<sup>6</sup>*

<sup>1</sup> EvalYouth- EvalYouth ECA Eastern Europe- Central Asia- South Caucasus, active member, Chisinau, Moldova Republic of

<sup>2</sup> EvalYouth Eastern Europe- Central Asia and South Caucasus, active member, Bishkek, Kyrgyzstan

<sup>3</sup> Eurasian Alliance of National Evaluation Associations, Chair, Yerevan, Armenia

<sup>4</sup> Ukrainian Evaluation Association- EvalYouth ECA, Board member- active member, Kyiv, Ukraine

<sup>5</sup> M&E community of Tajikistan- EvalYouth ECA, active member, Dushanbe, Tajikistan

<sup>6</sup> United Nations Population Fund, Director- Evaluation Office, New York, USA

This panel presents the overview of Implementation of EvalAgenda 2020 in the region Eastern Europe, Central Asia, and South Caucasus through the lens of Young and Emerging Evaluators (YEE) with the concentration on opportunities and challenges.

The main aim of the panel is to present the evaluation landscape of the region and the role of YEE, both actual and potential. The presentations are based on exploratory research results and secondary data analysis of investigations made by international and local non-governmental organizations. The panel makes a valuable contribution in an analysis of evaluation development in the region and provides insights into the situation of young and emerging evaluators, actual and potential role, achievements of YEE initiative towards building bridges for the effective start of evaluation career in the region.

In ECA region, there are not many capacity-building activities in evaluation, most of them concentrated around different donor programs or very narrow oriented. At the same time, there is a lack of local consultants in evaluation with both more profound knowledge of local context and understanding of global tendencies and achievements. Moreover, as a result of the low rate of English use, this region remains aside from global trends, and a poor activity on the development of evaluation systems remains. Therefore, the efforts of EvalYouth ECA are very important and valuable YEE opportunities and promotion of evaluation in the region in general.

EvalYouth ECA (Eastern Europe, Central Asia, and South Caucasus) was created in June 2017, and now it unites near 30 active YEE from Ukraine, Tajikistan, Armenia, Kyrgyzstan, and Georgia with more than 220 supporters in the mailing list. As a global network, EvalYouth and its regional chapter EvalYouth ECA has set out an ambitious goal to promote young and emerging evaluators, including young women, to become competent, experienced and well-networked professionals who contribute to evaluation capacity at national, regional and international levels.

The first presentation shows the overview of enabling environment for evaluation in ECA region with the particular role of Young and Emerging evaluators as drivers for the development. The second presentation concentrates on the development of Institutional capacities for capacity building in evaluation with emphasize on studying M&E programs and courses in the region. The third paper shows the importance of YEE in VOPEs development in the region and underlines the potential room for contribution. The fourth presentation makes the overview of individual capabilities for evaluation with strong emphasis on YEE opportunities and needs. Besides, it provides a mapping of YEEs in the region and potential spheres for growing of YEE initiative. The fifth presentation concentrates on achievements and challenges of EvalYouth ECA initiative in building inter-linkages for evaluation capacity development in the region. The last presentation underlines the importance of partnerships for YEE initiative and support if UNFPA for YEE program in the region.

**PN 56 EvalAgenda 2020 In the Region Eastern Europe, Central Asia & South Caucasus: Opportunities and Challenges for Young and Emerging Evaluators****PC 204 - Creating the Enabling Environment in ECA Region: the Role of YEE**

B. Zheenbekov<sup>1</sup>

<sup>1</sup> EvalYouth Eastern Europe- Central Asia and South Caucasus, active member, Bishkek, Kyrgyzstan

The first important dimension of building national evaluation capacity is creating an enabling environment for evaluation. This process involves a big number of stakeholders and combines multiple components. In order to realize a consistent enabling environment, the process should involve all social groups and spread into all social spheres. In 2017 people under 30 represents half of the world population, and those aged 15–24 sums around 1.2 billion. Although the age of professional career start often relies, later on, these numbers allow understanding the scale of the voice of young professionals providing that it will be properly heard.

Therefore, a proper involvement of YEE is a precondition for the successful development of an enabling environment for evaluation; they can be both supporters with active appreciation of evaluation within institutions (governmental institutions, academia, NGOs, the private sector, etc.), or/and evaluation advocators for the new implementation of evaluation systems.

There are external and internal influencers for developing an enabling environment. External facilitators are, mainly, international organizations, which provide some sort of aid and demand a strict process of monitoring and evaluation for project implementation. It could be marked as an obligation to conduct the evaluation. In many regions, especially it is true for Eastern Europe and Central Asia, international organizations involve external evaluators into their work, because of a higher level of competences and larger experience on required procedures. However, external evaluators cannot fully understand the context of the particular region, country; and, what is even more important; their activity is temporary and often does not lead to the well-grounded implementation of evaluation systems on the local levels. Empowering young professionals, and young and emerging evaluators, in particular, can bring new forces into the development of national evaluation capacity and create the broad pool of internal influencers, who day-by-day are building the base for enabling environment for evaluation.

**PN 56 EvalAgenda 2020 In the Region Eastern Europe, Central Asia & South Caucasus: Opportunities and Challenges for Young and Emerging Evaluators****PC 205 - YEEs as Drivers of Changes in ECA VOPEs**

T. Narchynska<sup>1</sup>

<sup>1</sup> Ukrainian Evaluation Association- EvaYouth ECA, Board member- active member, Kyiv, Ukraine

Cases of YEE involvement into VOPE developments/evaluation projects in ECA and Eastern-Central Europe; its analysis and recommendations.

National evaluation capacity highly depends on the existence of the relevant number of institutions, which support and use the evaluation. Countries with a high national evaluation capacity are characterized by proactive civil society, growing public demand for broader transparency and public accountability. In this process, YEE can play leading role as participants of grass root initiatives, creators of platforms for discussions and engagement of the broad audience. They can be also involved in governmental agencies, bringing innovative approaches to data usage and building the evidence-based policy.

Supporting professional organizations in evaluation can become a necessary impulse and strong advocacy tool for development of evaluation. High engagement of YEE can bring many benefits for improvement of such organizations. They often bring a fresh perspective on activity and procedures, YEE more freely uses technical tools, can bring innovative solutions according to recent trends, together with assuring future sustainability of the organization. More often they are educated in foreign environments, thus being connected to international professional networks and centers of expertise.

**PN 56 EvalAgenda 2020 In the Region Eastern Europe, Central Asia & South Caucasus: Opportunities and Challenges for Young and Emerging Evaluators****PC 206 - Development of Individual Capacities for Evaluation: Needs, Challenges, Opportunities**

*K. Rasulova*<sup>1</sup>

<sup>1</sup> M&E community of Tajikistan- EvalYouth ECA, active member, Dushanbe, Tajikistan

There is a great potential for the development of Youth Monitoring and Evaluation Movement in post-Soviet space. Monitoring and evaluation, especially in the countries of central Asia, is a relatively young direction that contributes to the effective implementation of national, social, economic, environmental programs and projects. But despite that, this movement is developing rapidly, and young professionals everywhere are needs to capacity building.

There are many prospects for young people with language skills (English) to develop and improve their M&E capacity. These include teaching programs abroad, the opportunity to participate in the program Eval Youth, and exchange of experience with colleagues, mentoring programs and others.

But along with this, there are both external and internal difficulties. External obstacles include access to information, the lack of educational programs for M&E in higher education and in view of evaluator's youth, the prejudiced attitude towards him of the older generation.

Internal obstacles include the lack of experience and skills in monitoring and evaluation. The process of monitoring and evaluation in practice, in addition to theoretical knowledge, requires knowledge of the local mentality, a complete understanding of the process being studied.

Often, young appraisers with only theoretical knowledge before entering the monitoring set a lot of indicators that as a result complicate the process and do not help to identify the "tree of the problem"

In Central Asia, people of the older generation are prejudiced against young Evaluators. Therefore, it is doubly difficult for young appraisers to explore social phenomena and processes.

We see the solution of these problems in the following:

- The introduction of the concept of monitoring and evaluation, its role and tasks in the process of teaching in schools and universities.
- Holding summer/winter schools for young evaluators with the participation of experienced specialists, exchange of experience and knowledge.
- Conducting practical training: a monitoring duet – an experienced evaluator + young evaluator
- Translation of the basic doctrines on M&E into Russian/national languages

**PN 56 EvalAgenda 2020 In the Region Eastern Europe, Central Asia & South Caucasus: Opportunities and Challenges for Young and Emerging Evaluators****PC 207 - Creating Strong Interlinkages: EvalYouth ECA as the Initiative of YEE in the Region and for YEE in the Region**

*A. Rishko-Porcescu*<sup>1</sup>

<sup>1</sup> EvalYouth- EvalYouth ECA Eastern Europe- Central Asia- South Caucasus, active member, Chisinau, Moldova Republic of

The strong inter-linkages among all dimensions facilitate the development of national evaluation capacity and create a platform for strong partnerships. The initiatives, which unite young and emerging evaluators with the aim to increase their professional capacity and confidence in career, at the same time, through empowering YEE they contribute to development evaluation capacity. EvalYouth made a great step towards it on the global and local level. Throughout the process, the support of YEE is very important, but not only by the senior evaluators, but peer-

to-peer interaction and experience sharing are providing the platform for professional start and growth, gaining confidence in the world of evaluation. EvalYouth is an initiative, which aimed to help YEE during this process. The initiative unites young and emerging evaluators with the aim to increase their professional capacity and confidence in career, at the same time, through empowering YEE. EvalYouth made a great step towards a common understanding on the global and local level.

EvalYouth ECA (Eastern Europe, Central Asia, and South Caucasus) was created in June 2017, and now it unites near 30 active YEE from Ukraine, Tajikistan, Armenia, Kyrgyzstan, and Georgia with more than 220 supporters in the mailing list. In ECA region, there are not many capacity-building activities in evaluation, most of them concentrated around different donor programs or very narrow oriented. At the same time, there is a lack of local consultants in evaluation with both more profound knowledge of local context and understanding of global tendencies and achievements. Moreover, as a result of the low rate of English use, this region remains aside from global trends, and a poor activity on the development of evaluation systems remains. Therefore, EvalYouth ECA starts working on opportunities for YEE and promotion of evaluation in the region.

### **PN 56 EvalAgenda 2020 In the Region Eastern Europe, Central Asia & South Caucasus: Opportunities and Challenges for Young and Emerging Evaluators**

## **PC 208 - Discussion of the Panel**

*A. Rishko-Porcescu<sup>1</sup>, M. Segone<sup>2</sup>, A. Harutyunyan<sup>3</sup>, B. Zheenbekov<sup>4</sup>, T. Narchynska<sup>5</sup>, K. Rasulova<sup>6</sup>*

<sup>1</sup> EvalYouth- EvalYouth ECA Eastern Europe- Central Asia- South Caucasus, active member, Chisinau, Moldova Republic of

<sup>2</sup> United Nations Population Fund, Director- Evaluation Office, New York, USA

<sup>3</sup> Eurasian Alliance of National Evaluation Associations, Chair, Yerevan, Armenia

<sup>4</sup> EvalYouth Eastern Europe- Central Asia and South Caucasus, active member, Bishkek, Kyrgyzstan

<sup>5</sup> Ukrainian Evaluation Association- EvalYouth ECA, Board member- active member, Kyiv, Ukraine

<sup>6</sup> M&E community of Tajikistan- EvalYouth ECA, active member, Dushanbe, Tajikistan

Marco Segone is a discussant in the panel. After each paper is presented, he will offer commentary on the papers presented and raise several important questions for the role of young and emerging evaluators in the implementation of EvalAgenda 2020, and evaluation development in general, at the global level and in the region Eastern Europe, Central Asia & South Caucasus.

Thursday, 4 October 2018  
17:30–19:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 57 Evaluation for Market Resilience: Methodological Challenges and Approaches**

## Evaluation for Market Resilience: Methodological Challenges and Approaches

S. Tenev<sup>1</sup>

<sup>1</sup> *The World Bank, Independent Evaluation Group, Senior Manager, Washington Dc, USA*

When functioning well, markets are key institutions for promoting prosperity, stability and inclusion. When they don't, they can exacerbate inequality and cause economic and social dislocations. Governments and non-government organizations alike increasingly put forward as part of their mandates the goal to establish and promote markets that are inclusive and resilient. Correspondingly, evaluation professionals increasingly have to grapple with the challenges of evaluating the merits of these efforts. This is particularly the case in economic development, where the development community has put a strong emphasis on the role of markets and the private sector in achieving the Sustainable Development Goals (SDGs). This panel will bring together practitioners in evaluation for economic development that deal with the challenges of evaluating the effects of development programs focused on strengthening or creating markets and making them more inclusive and resilient. Topics that will be covered and discussed include, but are not limited to: – Crisis prevention and crisis response: how to shift the balance? Recurrent crises of similar natures and often in similar contexts indicate limited ability to learn from experience. In addition to institutional and market characteristics are there behavioral patterns that underlie this trend? What is the role of evaluation in understanding the root causes of this phenomenon? Is there a need to incorporate behavioral change in evaluative approaches? – Economic Inclusion and Market Resilience – is there a need for integrated approach? While evaluators and researchers tend to look at market resilience and economic inclusion in isolation, growing evidence suggests that the two are closely related and that economic and social inclusion may be the most effective way to achieve stability and resilience. Do evaluators have the theoretical and empirical underpinnings to apply such an integrated approach to the assessment of development projects and programs? – Methodological challenges in evaluating development programs oriented at market creation. Achieving the SDGs require a higher level of ambition among development practitioners. Increasingly development organizations seek to achieve transformational changes through support to market creation. Some of the challenges that evaluation practitioners face in this regard include: what are the appropriate theory/ies of change for market creation? What criteria and standard of merit to use in assessing the effectiveness of these efforts? The panel will seek to advance the debate on these topics as well as help build skills and capacities among evaluation professionals in these areas based the experience of organizations that are leaders in the use of market-oriented approaches to development.

**PN 57 Evaluation for Market Resilience: Methodological Challenges and Approaches**

## PC 209 - Crisis Response and Crisis Prevention

B. Kolodkin<sup>1</sup>

<sup>1</sup> *EBRD, Deputy Chief Evaluator, London, United Kingdom*

Response to crises has been an important feature of the mandate of Multilateral Development Banks (MDBs). MDBs have spent vast amounts of money in a short period of time to help countries and regions recover from economic crises. The effectiveness of these crises responses has been also evaluated some of them in real time. What do these evaluations tell us about the effectiveness of crisis response by MDBs? And does crisis response/countercyclical lending lead to greater resilience and thus to crisis prevention?

**PN 57 Evaluation for Market Resilience: Methodological Challenges and Approaches****PC 210 - What Role can MDBs Play in Promoting Market Resilience?**

*T. Dimitrov*<sup>1</sup>

<sup>1</sup> Black Sea Trade Development Bank, Head of Evaluation, Thessaloniki, Greece

Among the multiple aspects of MDB's operations, what is the place and importance of their work on promoting resilient markets? What tools and approaches do the MDBs have to advance market resilience and how they succeed in going beyond being "banks" towards "knowledge banks," "change agents," and even "benchmark setters" in these efforts? What is the role of evaluation and the implications for evaluation methodology in this area?

**PN 57 Evaluation for Market Resilience: Methodological Challenges and Approaches****PC 211 - Inclusion and Resilience**

*M. Taylor-Dormond*<sup>1</sup>

<sup>1</sup> Asian Development Bank, Director General Evaluation, Manila, Philippines

The Multilateral Development Banks have multiple objectives. Among them the promotion of inclusion and resilience occupy a prominent position. What does evaluation evidence tell us about the relationship between inclusion and resilience? Do we find evidence of trade offs (as in the case of the sub-prime mortgage crisis in some developed countries) between the two and if yes, how are these trade-offs managed in development strategy and operations? What are the implications of these relationships between inclusion and resilience for evaluation methodology?

**PN 57 Evaluation for Market Resilience: Methodological Challenges and Approaches****PC 212 - Market Resilience in Private Sector Evaluation**

*F. Korfker*<sup>1</sup>

<sup>1</sup> Independent Consultant, Evaluation, Amsterdam, Netherlands

Private Sector Evaluations assess the effectiveness of development operations that use the market and the private sector as channels and mechanisms for impact. How does the methodology of private sector evaluations incorporate and reflect the peculiarities of the market as opposed to the non-market mechanism of allocation, which prevails in development assistance? How is market resilience reflected in the criteria and standards of merit employed by private sector evaluation?

Thursday, 4 October 2018  
17:30–19:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 58 The Power Of Diversity And Inclusiveness (DI): A Path To Peace And Prosperity**

## **The Power of Diversity and Inclusiveness (DI): A Path to Peace and Prosperity**

*S. Premakanthan*<sup>1</sup>

<sup>1</sup> *Symbiotic International Consulting Services SICS, Consulting, Ottawa- Ontario, Canada*

3 Panelist will discuss: Panelist Leader Sandiran (Sandi) Premakanthan will introduce the topic and discuss global research findings Diversity is any dimension that can be used to differentiate groups and people from one another. It means respect for, and appreciation of, differences in ethnicity, gender, age, national origin, disability, sexual orientation, education and religion. Inclusiveness is a state of being valued, respected, involved and supported. The McKinsey research findings support the view that companies with more diverse workforces perform better financially creates shareholder wealth. Diversity's dividend can be viewed as EBIT – earnings before interest and taxes. One of the key highlights of the McKinsey report is that "diversity is a competitive differentiator". I think the trend worldwide is showing an appetite for implementing well designed DI programs led by a diverse and inclusive management hierarchy. In my view, organizations who embrace the true meaning of the definitions of DI will prosper and provide a path to peace and harmony in the workplace for all to contribute and realize their potential as valued members at work and society. Outcome Management and evaluation professionals should take note of the importance and strength of DI in developing and implementing purpose-driven evaluation management strategies to unleash the power of DI in our quest for excellence in evaluation and oversight functions. Organizations could leverage diversity and inclusion for their growth and success by providing professional management advice and solutions that reflect the diversity of clients. John Flanders (Panelist #2) will discuss: The working session will present how Canada's Prime Minister is leading the way on the world stage and Canada's progress on DI. A demographic snapshot of Canada's diversity and measures adopted by the Canadian government: Values & Ethics Code and Canada's Standard for Psychological Health and Safety in the Workplace, Duty to accommodate, New Joint Union/ Management Task Force and Multi-Year Employment Equity Plan. Atiq Rehman (Panelist #3) will discuss Participants will examine and engage in discussions on DI since South Africa became a democracy two decades ago. Government and unions have been at the forefront of championing employment equity in the workplace. The discussion will focus on the progress made, for example, the South African standard on employment equity, diversity and inclusion and the barriers to implementing strategies to harness the power and strength of DI. He will also discuss DI case study in Malaysia.

Thursday, 4 October 2018  
17:30–19:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 59 Evaluating Unintended Effects and Adapt in Countries Affected by Fragility, Conflict and Violence**

## **Evaluating Unintended Effects and Adapt in Countries Affected by Fragility, Conflict and Violence**

*H. Hassnain*<sup>1</sup>

<sup>1</sup> *Y Care International, Impact- Monitoring- Evaluation and Learning - Fragility- Conflict and Violence, London, United Kingdom*

Some of the most critical factors in achieving SDGs in countries affected by conflict are by efficiently understanding conflict dynamics to minimise harm and establishing effective partnerships to gather and analyse data with state, civil society, academia and the private sector. One key reason for a lapse in this is the dearth of effective monitoring and evaluation practices and trust between different actors operating in volatile contexts due to a variety of reasons. This discussion will mainly dwell upon the conference theme of M&E and the SDGs. The panellists will share their practical experiences about unintended effects, political agendas behind aids and real-time stories from all around the world.

Besides all other complications, evaluation in situations of conflict entails some serious human risks especially for the locals. This calls for a serious discussion on the available ethical protocols of evaluations in such environments. The data reported in the Aid Workers Security Database, according to which 80% of the estimated 208 aid workers killed, kidnapped or seriously wounded worldwide in 2016 were local people, is a clear pointer to this need.

Building on Michael Bamberger's research on the evaluation of unintended effects broadly defined, to focus more specifically on the less developed practice of assessing unintended escalations of violence or inter-group tensions. We will also discuss monitoring, because if negative effects are identified during implementation, then programmes can be adapted quickly to minimise harm and maximise benefits. Michael Patton affirmed that there is no one best way to conduct an evaluation and its depends, among other factors, on recognizing situational practices competencies in an evaluator.

This panel discussion will take stock of the efforts made and lessons learned in countries affected by social violence and conflict and will specifically look into conflict-sensitive practice that has never made it into the mainstream." Thania Paffenholz contribution in 'Conflict Sensitivity: Taking it to the Next Level'.

This panel will bring people together in thinking about how to measure and assess the unintended negative effects in conflict with a special focus on Sustainable Development Goal (SDG) 16, conflict sensitivity and continuous adaptation. Interested participants will be invited to continue post-conference communication on these issues.

Some of the questions that we will seek to answer together are:

- Impact evaluations for peace building, what are the challenging preconceptions?
- How can conflict sensitivity efforts be effectively monitored and evaluated, in ways that promote robust accountability at all levels, from providers to local communities to organizational headquarters and donors?
- How to measure the unintended negative effects in conflict, fragility and marginalization?
- What is the role of monitoring (including third party monitoring) in conflict and how can it be used to establish a flexible M&E system.
- What personal competencies should an evaluator have in addition to the techniques and approaches?

**PN 59 Evaluating Unintended Effects and Adapt in Countries Affected by Fragility, Conflict and Violence****PC 289 - Evaluating Unintended Effects and Adapt in Countries Affected by Fragility, Conflict and Violence**

H. Hassnain<sup>1</sup>

<sup>1</sup> Y Care International, Impact- Monitoring- Evaluation and Learning - Fragility- Conflict and Violence, London, United Kingdom

Contribution of Hur Hassnain, the chair.

**PN 59 Evaluating Unintended Effects and Adapt in Countries Affected by Fragility, Conflict and Violence****PC 220 - The Use of Art in Reconciliation and Memory Processes**

I.L. Aronsson<sup>1</sup>

<sup>1</sup> Uppsala University- Uppsala- Sweden, Resettlement- and heritage- memories and reconciliation in post-conflict areas, Uppsala, Sweden

This paper reflects on the enactments of culture, and how art projects can be used in reconciliation and memory processes to impact on societal transition after conflict. Art is viewed as a cultural-historical agent, and not as an isolated phenomenon. Art projects are important for peace processes, but we still struggle to understand how and why they help in the contemplation and healing of loss, trauma, mourning and memory. Furthermore, how does this subjective, and assumed collective experience connect to and manifest itself in the necessary political change (political reconciliation)? In this explorative paper, I elaborate on these crucial dimensions of transformation by presenting and analyzing selected art projects from Colombia.

**PN 59 Evaluating Unintended Effects and Adapt in Countries Affected by Fragility, Conflict and Violence****PC 221 - Evaluating Unintended Effects**

S. Tamondong<sup>1</sup>

<sup>1</sup> International Development Evaluation Association, Development Evaluation, Paris, France

In evaluating unintended effects and adaptation in countries affected by fragility and conflict, it is important to give attention to women. Women are particularly marginalized and under represented in fragile states and in post conflict societies. Around half of the world's poor live in fragile states and majority of them are women. As fragile states make people poor, fragility make people more at risk of poverty, making it harder for them to escape poverty. Women living in fragile states are even more marginalized and vulnerable to poverty. Six out of ten of the world's poorest are women. (UNDP Gender and Poverty Reduction report). The risks of poverty are even greater for women living in fragile states. Fragile states reinforce gender inequality. Fragile states turn to a patriarchal world view, traditional, religious and customary laws due to state weaknesses and ineffective institutions. Fragile states lack democratic accountability, which is a challenge to women. One in 3 women worldwide are victims of domestic or sexual violence in their lifetime, and women in fragile states are particularly at risk of being victims of violence. There is a higher level of societal violence and acceptance of violence in fragile and conflict areas. (WHO Report 2013 Violence against Women..." Fact Sheet 239.

If the need to promote women representation is not structurally addressed, early on, in the rebuilding of nations, as a matter of priority, rather than as add on, women's contribution in political and social processes, will be severely undermined. It would be a lost opportunity.

Women champion change beneficial to everyone – they raise social and practical issues, related to health, children, education and livelihood, water use and sanitation, issues women encounter in daily life- which men in power may not see easily, or at all. Strategies to address and improve accountability and governance must build women's participation. Women can be agents of change promoting social justice. Ways to involve women in formal and informal negotiations in transitional democracy and political settlements must be found. Women must be allowed to define and experience justice. Psychosocial support to women and girls in the aftermath of conflict is essential for lasting peace and development. There is a need to unlock gender equality in fragile states and conflict affected situations. Thus, in evaluating unintended effects and adaptation in countries affected by fragility and conflict, the role of women is imperative.

**PN 59 Evaluating Unintended Effects and Adapt in Countries Affected by Fragility, Conflict and Violence**

**PC 222 - Addressing M&E Challenges, Technical and Personal Evaluators Competences in Complex and Conflict Contexts**

*E. Rotondo*<sup>1</sup>

<sup>1</sup> ReLAC, Latin American Evaluation, Peru, Peru

The presentation will address the conceptual and methodological challenges to the M & E of programs in changing contexts of conflict and social violence that affect women, children and adolescents in particular. How to measure and verify expected and unanticipated changes in these contexts? How to approach the sensitivity and needs of affected populations? It starts from the definition of evaluation as a trans-discipline and public good that requires pluralism in its approaches to address complexity and capture different perspectives and needs. Likewise, it requires continuous monitoring systems focused on the people that allow to capture expected and unexpected results that favor social learning, not only institutional. The Latin American region has best practices and lessons learned from these M & E systems that may be useful for other regions. Also, Hur Hassnain will share and reflect on his experience working globally in countries affected by conflict and fragility. The presentation will also discuss a proposal of competencies necessary for internal and external evaluators to facilitate processes, learn and improve interventions in complex and conflict contexts.

This presentation will bring people together in thinking about how to measure and assess the unintended negative effects in conflict with a special focus on Sustainable Development Goal (SDG) 16, conflict sensitivity and continuous adaptation. Interested participants will be invited to continue post-conference communication on these issues.

Thursday, 4 October 2018  
17:30–19:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 60 Psychological Approaches in Strengthening Resilience of Sustainable Development Programmes From An M&E Perspective**

**Psychological Approaches in Strengthening Resilience of Sustainable Development Programmes from An M&E Perspective**

K. Odhiambo<sup>1</sup>

<sup>1</sup> University of Nairobi, Psychology, Nairobi, Kenya

**Rationale:** The Arising ecological problems of Sustainable Development (SD) based on ecosystem restoration such as Climate Change(CC), has resulted in calls for “Psychological Behaviour Science Corra-Verdugo (2006) beyond the frame of social, economic and political. Interest is in “human nature” within a “Psychological Behavior Science” predispositions thus the critical inter-link between adaptation of sustainable lifestyles and human livelihood. That the way humans behave has an effect on environment. Thus, Behaviour Science is needed to the fulfillment of sustainable purpose the focus of the world today towards sustainable livelihoods. United Nations (UN:2016) states, it is time for programmes to embrace behaviour attributes in any development.

**Objective:**

- a) To develop knowledge on the area of behaviour science as perceived and to share
- b) Ultimately, product will be a model to guide evaluation practice. The students will continue to be part of this on-going process
- c) To mentor the young evaluators within the evaluation fraternity to ensure skilled young evaluators on M&E and behaviour science

These are the-4 panelists who are Post-Graduate students of the class of 2017 the course being “Programme Evaluation and Development” with the LEAD Panelist as the Instructor

**Narrative and justification:** Sustainable Development can only reach it's fulfilment with human elements incorporated thus behaviour science. For the evaluation fraternity, this brings to bear adjustment into methods and approaches as well as logical theory based frameworks that are defined by sustenance and evaluation that should be re-thought through with new indicators, measures and impact forms determined. The discourse around psychology is new resulting in the need to guide M&E practitioners and so the panel presentation of sustainable psychologists with evaluation background.

Thus:

“... Through human way of life, once the resources are depleted, human beings turn to natural resources starting a long chain of concerns on \_ global warming, species and habitat depletion, chemical pollution, political conflicts...”

“... The result is Ecological effects of climate change, melting glaciers, green house gases, extinction of plant life...”

“... Impacts of ecology-climate that arise pauses threats to human health as it relates to physiological, psychological behavior perspectives, stressors on humans, and effects on choice s and decisions made ...”

“... In terms of coping appraisal as rivers dry up there is need for threat appraisal... what people think they could do or should do and change of behaviour needed

“ ... M& E needs to be integrated and should be integral to this to measure behaviour...”

The 5-panelists will cover:

- 1<sup>st</sup> Overview by CHAIR\_ Dr. Karen T Odhiambo
- 2<sup>nd</sup> Behaviour Science\_ the concept: Dianah Mwangi
- 3<sup>rd</sup> Behaviour models \_sustainable in nature: Doreen Kabiru

4<sup>th</sup> Behaviour science and Climate Change: Jedida Sakwa

5<sup>th</sup> Monitoring and Evaluation applications and Behaviour Science: Brenda

Behaviour arising leads to participation in environment as civic action at Policy, Administrative, community and household levels' \_ beyond driving individuals and communities to make decisions and act in response to stressors created as a result of the challenging context.

This is work in progress.

**PN 60 Psychological Approaches in Strengthening Resilience of Sustainable Development Programmes From An M&E Perspective**

**PC 223 - Psychological Approaches in Strengthening Resilience of Sustainable Development Programmes from an M&E Perspective**

K. Odhiambo<sup>1</sup>

<sup>1</sup> University of Nairobi, Psychology, Nairobi, Kenya

**Rationale:** The Arising ecological problems of Sustainable Development (SD) based on ecosystem restoration such as Climate Change(CC), has resulted in new perspectives beyond the frame of social, economic and political to approaches related to human nature" within a "Psychological Behavior Science" predispositions Corra-Verdugo (2006). Thus the critical inter-link between adaptation of sustainable lifestyles and human livelihood. That humans behave has a direct effect on environment is a given case of interest to evaluators. Therefore Behaviour Science applications is needed to complete the fulfillment of sustainable purpose the focus of the world today towards sustainable livelihoods. United Nations(UN:2016) states, it is time for programmes to embrace behaviour attributes in any development, so does the American Evaluation Association[APA].

**Objective:**

- a) To develop knowledge on the area of behaviour science as perceived and to share
- b) To mentor the young evaluators within the evaluation fraternity to ensure skilled young evaluators on M&E and behaviour science. These are the-4 panelists who are Post-Graduate students of the class of 2017 the course being "Programme Evaluation and Development" with the LEAD Panelist as the Instructor
- c) Ultimately, product will be a model to guide evaluation practice.

The students will continue to be part of this on-going process

**Narrative and justification:** Sustainable Development can only reach it's fulfillment with human elements incorporated thus behaviour science. For the evaluation fraternity, this brings to bear adjustment into methods and approaches as well as logical theory based frameworks that are defined by sustenance and evaluation that should be re-thought through with new indicators, measures and impact forms determined. The discourse around psychology is new resulting in the need to guide M&E practitioners and so the panel presentation of sustainable psychologists with evaluation background.

Thus:

"... Through human way of life, once the resources are depleted, human beings turn to natural resources which starts a long chain of concerns regarding \_ global warming, species and habit depletion, chemical pollution, political conflict..."

"... The above results in Ecological effects of climate change, melting glaciers, green house gases, extinction of plant life..."

"... whereby impacts of ecology-climate that arise pauses a threat to human health as it relates to physiological, psychological behavior perspectives, stressors on humans, and effects on choices and decisions made ...."

"... In terms of copying appraisal as rivers dry up there is need for threat appraisal... what people think they could do or should do and change of behaviour needed

" ... M& E needs to be integrated into this process..."

The 5-panelists will cover:

- 1<sup>st</sup> Overview\_ CHAIR
- 2<sup>nd</sup> Behaviour Science\_concept;
- 3<sup>rd</sup> Behaviour models \_sustainable in nature;
- 4<sup>th</sup> Behaviour science of \_ Climate Change;
- 5<sup>th</sup> Monitoring and Evaluation applications of sustiance

Behaviour arising leads to participation in the environment as civic action at Policy, Administrative, community and household levels'\_ beyond driving individuals and communities to make decisions and act in response to stressors created as a result of the challenging context.

The product intended is presented as work in progress.

### **PN 60 Psychological Approaches in Strengthening Resilience of Sustainable Development Programmes From An M&E Perspective**

## **PC 224 - Conceptualising Behaviour Science Within Sustainable Development[SD] Perspective**

*D. Mwangi<sup>1</sup>*

<sup>1</sup> *University of Nairobi, Psychology, Nairobi, Kenya*

Rational Human behaviour and relation to ecology is important more so the way people behave and respond more so as a result of the effects that arise from mitigations such as Climate Change. In SDGs this result in mitigation and adaptive survival traits. This could be negative or positive depending on the extent to which the behaviour response leads to depletion of the environment. Objective: Is to introduce evaluators to the behaviour perspectives and their link to sustainable circumstances Narrative and Justification: It is important to understand the link between behavior and environmental factors and effects. The issue is there is not much knowledge in this area as it is new and the world is only now turning to behaviour science to enhance development response more so sustainability. This section will attempt to conceptualise and introduce the audience to this link. There is need to determine review the arising psychological adaptive measures such as appraisal of the situation; cognitive analysis of thought processes; disengagement measures, defensive responses, and emotional regulation. There is also issues such as proximity, exposure, sense making, social construction, and risk as well as collective efficacy This is based on the fact that behaviour science perspective is indispensable for increasing understanding of the drivers of SDGs. Behaviour perspectives that arise consist of intra-individual parameters and processes. First behaviour will be defined, behaviour being a psychological attribute and generally referred to as the "the science of human behaviour and mental processes." Thinking and feeling is included here. It has it's roots\_ perception, attitude, values and beliefs. Each attribute results in how we response during on-going environmental experience. This could result in conditional habits or behaviourism; learned behaviour or social; thinking and acting by reason or cognition; and taking responsibility for the actions we decide on or humanistic. The term survival comes to mind here. The behaviour aspects that arise addresses the following parameters: appraising situations,, affective response or emotions, cognition or analysis through reasoning, disengaging or separating as well as defensiveness or self justifying. This brings with it collective sense making or shared and social comparison or communality. The behaviour scenario that arise are many regarding peoples environment. The presenter will show this by exhibiting what happens. For example: "...there is at such a time threat appraisal arises \_ what the affected should do and if it will make a difference..." "... There is also Risk Perception \_ which arises from dread, uncertainty catastrophe, prior viscious experience and psychological distance..." The presenter will also show that behaviour is complex and be understood from a theoretical basis and fundamentals of programme logic and form. This paper is part of the PANEL Presentation of The class of 2017 who took a course on "Programme Evaluation and Development" \_ Taught and guided by Dr. Karen Odhiambo\_ University of Nairobi. This was from the classroom Task on SDGs and M&E.

**PN 60 Psychological Approaches in Strengthening Resilience of Sustainable Development Programmes From An M&E Perspective****PC 225 - Behaviour Science Principles Models and Environmental Sustainable Development**

D.K. Kabiru<sup>1</sup>

<sup>1</sup> University of Nairobi, Psychology, Nairobi, Kenya

**Rationale:** Many behaviour models arise however it is important to provide a model or frame of thought derived from the logic of sustainable circumstances that is situation specific. Whether specific or general, the schemas that arise are similar. Approaches that follow Climate Change for example has followed a profile that is individual and communal based. One has to determine the theory and logic that underlies sustainable programmes. The frame of thought envisaged will be based on the underlying aspects of the commonalities that arise in psychologically. The psychological models that arise are related to environmental stress; adaptation; stress and coping models; protection motivation theory; and health well being.

**Objectives:** a) To guide in determining frame of thought on psychological models and SD b) The schemas that arise when considering a programme on SD

**Narrative and Justification:** In order to ensure that new behaviour approaches are adaptable and within reach in terms of knowledge and form, there is need to provide guidance to practitioners. Thus an evolving guide and frame of thought is desirable. This is important as this is an area that is new and is still being explored. Behaviour perspectives are general and complex and cuts across many fields such as organisations, individuals, communities, policies, politics etc. There is need to harmonise what behaviour science is and to come up with a comprehensive approach. This is the basis of this presentation. The approach is more to provide a guideline other than a set of rules and regulations. Specific elements of the model/theory are related to: those that include direct and indirect experiences with threats and impacts of Climate change; appraisals of impacts related to resources as threat appraisals; Appraisals of possible responses or copying appraisals; emotional based; and motivational responses. The behaviour dimensions that arise are related to but not exclusive to: Resilience, stressors, empowerment, self-determination, adaptation, and well-being. There is also the organizational that includes policy design and levels of sustenance governance This paper is part of the PANEL Presentation of The class of 2017 who took a course on "Programme Evaluation and Development" \_ and was taught by the LEAD panelist \_ Dr. Karen Odhiambo \_ University of Nairobi. This was from the classroom Task on SDGs and M&E.

**PN 60 Psychological Approaches in Strengthening Resilience of Sustainable Development Programmes From An M&E Perspective****PC 226 - Behaviour Science Aspect of Mitigating Factors in Climate Change: Towards Sustainable Communities**

J. Sakwa<sup>1</sup>

<sup>1</sup> University of Nairobi, Psychology, Nairobi, Kenya

**Rationale:** Interest arises from the fact that mitigating factors such as Climate Change(CC) as perceived regarding issues of global warming include psychological predisposition that would best drive the adoption of sustainable lifestyle. They help bring out existence, routines, means to, and regimes of behaviour which are important aspects of human livelihood and well being. They are important as they affect and influence behaviour that is expected to lead to empowerment of humans in ecosystem restoration.

**Objectives:** a) To determine behavior perspectives that mitigate factors such as Climate Change b) Behavior influence that lead to EMPOWERMENT within ones locale of sustainable development

**Narrative and Justification:** Given environmental concern and strategy of global warming, species and habitat, change in temperature, chemical pollution and political conflicts \_ and there after the ecological effects of climate change, endangered species, melting glaciers, pollutants, green house gases, access to clean water, conflicts and war \_ the effects on human relations that arise result in psychological, physiological, resilience and effects on choices made on the environment. Thus looking at the human person local level of existence, within their locale of development, psychological perspectives arise. This has an effect on the future state of the ecosystem. This is inbuilt into scenarios such as when the ecosystem is not balanced, when people are poor they turn to the environment. Thus stress on environment results on human stress ecologically and finally depletion of the environment. The presenter will show the link between CC factors and psychological perspective using evaluation link. The presenter will also show the nature of behaviour science that arise that need to be considered in determining CC programme as an aspect of SDGs locally. Given that humans contribute to ecosystem destruction, this is intimately related to sustainability and behaviour. CC is therefore a central concern of psychology (and other behaviour sciences such as organisations and economics). Thus the presenter will eliminate the behaviour aspects of CC by: the behaviour psychological responses to CC impact and possible barriers, consider behaviour focused intervention strategies, human dimensions, adaptation arising and it's behaviour sequel etc. Further it will show hoe evaluation plays a role in guiding the theory of logic in creating synergies and partnerships. This paper is part of the PANEL Presentation of The class of 2016 who took a course on "Programme Evaluation and Development"\_ Taught by the LEAD panelist, Dr. Karen Odhiambo\_ University of Nairobi. This was from the classroom Task on SDGs and M&E.

#### **PN 60 Psychological Approaches in Strengthening Resilience of Sustainable Development Programmes From An M&E Perspective**

### **PC 227 - Monitoring and Evaluation Aspects of Behaviour and It's Link to Sustainable Circumstances**

B. Gutu<sup>1</sup>

<sup>1</sup> University of Nairobi, Psychology, Nairobi, Kenya

**Rationale:** A paradigm shift towards fulfilling the goal of SDGs has been recognized and appreciated. American Psychology Association (APA:2014) states that this is the moment for psychologists and evaluators such as this PANEL should recognize this paradigm shift and to recognize their important role. This is because, development can only become to its fulfilment with human elements incorporated into it, such as behaviour. For evaluators this bring to bear adjustment into methods and approaches as well as logical theory based frame work of programmes that should be revised to adapt to behaviour aspects of sustainability and indicators identified and applied and measures and impacted determined. The discourse around this is new and so the need to share what this means to M&E practitioners is important.

**Objective:** a) To demonstrate the paradigm shift within an M&E perspective and Behaviour Science b) The link of M&E with Sustenance c) Programme Theory and Logic as regards SDG and Behaviour Science

**Narrative and Justification:** This comes as the evaluation fraternity is still evolving much as it has made great gains. The challenge is two fold in that evaluation mechanisms are still evolving much as a lot of gains has been made. The issue of sustainable development in regard to SDGs and even CC has resulted in many schemas and sequels of levels of analysis and measures needed to adequately frame SDGs within behaviour science perspective. There is need to come up with a guide that will guide not only understanding and conceptualization of such a process but also levels that arise within a programme of such nature. Thus aspects of planning, interventions enhancing livelihoods, community adaptations are yet to be constituted into knowledge that can address the new dispensation. Further the attributes that encompass disaster because of multiple impacts at any one time further complicates this process not forgetting individual and community as well as house hold levels of psychological process. There is need to determine review the arising psychological adaptive measures such as appraisal of the situation; cognitive analysis of thought processes; disengagement measures, defensive responses, and emotional regulation. There is also issues such as proximity, exposure, sence making, social

construction, and risk as well as collective efficacy. Of importance is their measure and indicators that are generally subjective. Such things as “actions does not mean intent” become important. Thus motivation becomes a key issue as well as rewards and punishment forms. The logic in formulating the situation arising will be presented to include aspects of M&E to guide and support practice. This paper is part of the PANEL Presentation of The class of 2017 who took a course on “Programme Evaluation and Development”

The Lecturer is the LEAD panelist, Dr. Karen Odhiambo, University of Nairobi. This was from the classroom Task on SDGs and M&E.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 61 Using Big Data to Evaluate an Experimental Program Addressing Poverty, Climate Change and Resilience among Small Farmers in SE Asia**

## Using Big Data to Evaluate an Experimental Program Addressing Poverty, Climate Change and Resilience among Small Farmers in SE Asia

*M. Bamberger*<sup>1</sup>

<sup>1</sup> *Independent consultant, No university affiliation, Beaverton, USA*

This panel reviews lessons from the application of digital technology in addressing poverty and climate change among small rice producers in South East Asia, and considers how program generated big data could be used to develop an evaluation model, to assess the impacts of this program on small-holder farmers. Lessons will also be drawn from other digitally driven poverty reduction programs promoted by The Development CAFÉ with respect to the current debates on the opportunities and challenges for integrating big data into the evaluation of development programs.

In addition to widespread poverty, smallholder rice cultivation is a major source of carbon emissions. The goal of the Stick Rice Digital Platform is to combine poverty reduction with climate change through training, supply chain traceability, the use of block chains to promote participatory guarantee systems, payment and credit services, and carbon estimates, verification and financial incentives. The project will be experimenting with and evaluating different kinds of incentives to reduce emissions. All stages of the program are supported by innovative uses of digital technology. The goal is to reach 1 million small farmers, so the quantity and variety of data generated will illustrate the potential applications of big data and data analytics in the program evaluation. The aim of the presentation is to present a case study and also take inputs from evaluators to develop methodologies in light of these emerging technologies, that are here to stay and how development evaluators can guide and harness them.

The first presentation gives an overview of the project and how it will apply digital technology to address poverty, climate change and resilience. It will review the kinds of big data that will be collected; and will also present the program theory of change, identifying the intended outputs and outcomes and the processes through which they are to be achieved. The presentation will also draw lessons on the opportunities and challenges for integrating digital technology into development projects targeted at very poor and dispersed populations. The presentation will also explain how the project will address resilience. The evaluation design is still under discussion, and the second presentation will draw on recent discussions in the data science literature to explore some of the new evaluation design options that are being considered, and how they would integrate data science and traditional evaluation approaches. The possibility of using some of the new experimental and quasi-experimental design options that data analytics are opening-up, will be considered. At the same time the importance of a mixed methods approach, combining big data and conventional evaluation approaches will be emphasized. If time permits, the importance of a complexity-responsive evaluation framework will also be introduced. The discussion will draw lessons from this project within the broader context of some of the current debates and approaches on the applications of big data in development and the integration of data science and evaluation.

**PN 61 Using Big Data to Evaluate an Experimental Program Addressing Poverty, Climate Change and Resilience among Small Farmers in SE Asia****PC 228 - The Sticky Rice Platform for Addressing Poverty, Climate Change and Resilience Among Small Farmers in Se Asia**

V. Gandhi<sup>1</sup>

<sup>1</sup> *The Development Cafe, Research- Evaluation- Innovation and ICT Unit, Jakarta, Indonesia*

The first presentation gives an overview of the project and how it will apply digital technology to address poverty, climate change and resilience. It will review the innovative big data collection methods and the kinds of data that will be collected. It will also present the program theory of change, identifying the intended outputs and outcomes and the processes through which these will be achieved. Some of the critical hypotheses and assumptions will be identified and how they are to be tested. The presentation will also draw lessons on the opportunities and challenges for integrating digital technology into programs targeted for very poor and remote populations. This will also draw lessons on the many practical, logistical and organizational challenges in introducing and maintaining these kinds of technologies. The perspectives of different stakeholders, including funders, implementing agencies and the small farmers will be reviewed. One of the important new technologies that is being introduced is block chains. Some of the potential benefits for increasing data security, increasing the confidence of different stakeholders in the information systems and also increasing the quality of the information and financial systems and the ability for farmers and other stakeholders to increase their control over the kinds of information that can be shared and protected. The presentation will also explain the different ways that the project will promote resilience and address climate change.

**PN 61 Using Big Data to Evaluate an Experimental Program Addressing Poverty, Climate Change and Resilience among Small Farmers in SE Asia****PC 229 - Exploring Research Design Options for the Sticky Rice Program to Integrate Big Data and Conventional Evaluation Designs: Opportunities and Challenges**

M. Bamberger<sup>1</sup>

<sup>1</sup> *Independent consultant, No university affiliation, Beaverton, USA*

The second presentation will describe the different approaches that are being explored for the evaluation of individual projects and for the overall evaluation of the multi-project/multi-location Sticky Rice project. The operational aspects of the data collection and analysis are still being put into operation and tested. Consequently, several possible options for the evaluation design are still being explored, but an initial design (or designs) will have been identified by the time of the conference in October. These will draw on the still relatively small, but rapidly expanding literature on data science-based evaluation designs. Examples of the approaches that are being explored and that will be discussed include: running big data-based experiments (e.g. Salganik 2018 "Social research bit-by-bit in the digital age"), social media analysis (including social network analysis and the extensive body of research based on Twitter and Facebook); using satellite images and remote sensors to strengthen propensity score matching in comparison group designs (for example the work of the Global Environment Facility); and predictive analytics. We will also draw on recent discussions on the World Bank impact evaluation blog on potential applications of machine learning and AI to strengthen different elements of impact evaluation design such as: broadening the range of outcomes that can be measured, improving the targeting and design of the treatment, better estimates of treatment heterogeneity, and controlling for confounders. These are just examples of approaches being explored. A central premise of the approach to evaluation will be the importance of using a mixed methods approach that combines big data with conventional evaluation designs. The presentation will also discuss opportunities and challenges for the integration of big data and conventional approaches to evaluation. While the opportunities and benefits of an integrated approach are evident, there are several methodological, organizational, logistical, political and even philosophical that must be addressed.

**PN 61 Using Big Data to Evaluate an Experimental Program Addressing Poverty, Climate Change and Resilience among Small Farmers in SE Asia**

## **PC 230 - Discussant**

*E. Raimondo*<sup>1</sup>

<sup>1</sup> World Bank, Washington DC, USA

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 62 Participatory Approaches to Resilience Measurement: Who are we Listening to?**

## Participatory Approaches to Resilience Measurement: Who are we Listening to?

*J. Cekanova*<sup>1</sup>

<sup>1</sup> *ValuingVoices at CEKAN CONSULTING LLC, President/ Catalyst, Prague, Czechia*

Resilience has established itself as fixture in the humanitarian and development industries. Given the increased funding devoted to interventions packaged as 'resilience-strengthening', measurement and overall accountability of these interventions to the participants whom we ostensibly serve is growing. The panel will discuss how participatory approaches engage local men and women.

There are many definitions and even more manifestations of Resilience. Generally, resilience can be defined as "the capacity of individuals, families, communities, systems, and institutions to anticipate, withstand and/or judiciously engage with catastrophic events and/or experiences; actively making meaning with the goal of maintaining normal function without fundamental loss of identity". (Almedom and Turnwine, 2008). Resilience is no longer a buzzword, but has established itself as fixture in the humanitarian and development industries. Given the increased funding devoted to interventions packaged as 'resilience-strengthening', resilience measurement and the overall accountability of these interventions to the participants whom we ostensibly serve is growing. Some of these measurement systems use more quantitative methods that yield objective comparison over time of 'who, what, where and how much' versus more qualitative, even more subjective measures but richer, context-specific 'how, with whom and why'.

This session will examine considerations for evaluating resilience-strengthening from participatory and gender lenses as well as from changing what theories of change look like from our participants' perspectives. The panel will discuss how participatory approaches engage local men and women in M&E, so that programs can more effectively foster our participants' own capacity for resilience. How? A start is supporting a process that fosters their voices to inform and define what capacity towards positive resilience looks like, how to elicit their views about local realities and needs and how to shape more resilient programs while balancing input to ensure evaluators are not misled. The panel will discuss how participatory approaches engage local men and women.

**Presentations:** Scott Chaplowe – While there is a value to standardized, quantitative approaches for impact and attribution analysis, this presentation will highlight the limitations of such approaches for the measurement and evaluation of resilience-strengthening interventions, and will highlight the importance of context-specific, participatory methods that allow resilience to be defined and measured according to community realities and needs.

Sanjeev Sridharan- A focus on gender and inequalities may need to consider intersections between gender, class, race, education, caste, etc. and the power relations that drive such hierarchies. How can participatory M&E and analysis illuminate such intersections and hierarchies.

Isabella Jean, CDA's Listening Project reached 6,000 people in 20 countries and listened to people experiencing chronic vulnerability, slow-onset and rapid-onset disasters and conflicts, noting striking consistency in their descriptions of how aid organizations do not sufficiently understand or consider the system of relationships they are stepping into which is the cornerstone of the Do No Harm approach. Recent commitments to test flexible funding mechanisms also require an equal attention to inclusive and broad community participation in shaping priorities and engaging people in decision-making affecting their lives.

Colin McQuistan, Practical Action UK piloted a participatory resilience measurement tool in over 100 communities in 11 countries over the last 3 years. The results from the measurements vary hence very different interventions have been planned and implemented. The differences in approaches expose different depths of engagement and early indications are that the best outcomes come from those engagements that have reached the weakest links across social, cultural, economic and environmental dimensions. The presentation will highlight some of these human stories that demonstrate very different perceptions of resilience.

**PN 62 Participatory Approaches to Resilience Measurement: Who are we Listening to?**

## **PC 231 - Rethinking Evaluation Methods and Methodologies**

J. Cekanova<sup>1</sup>

<sup>1</sup> *ValuingVoices at CEKAN CONSULTING LLC, President/ Catalyst, Prague, Czechia*

Jindra Cekan, PhD – Chair, posing provocative questions about who defines participation, what M&E tools are used, how lessons are shared, how do we know what 'truth' is in participatory evaluation, and how things change.

**PN 62 Participatory Approaches to Resilience Measurement: Who are we Listening to?**

## **PC 232 - Participatory Approaches to Resilience Measurement: Who are we Listening to? Context Matters**

S. Chaplowe<sup>1</sup>

<sup>1</sup> *Children's Investment Fund Foundation, Director of Evidence- Measurement and Evaluation Climate, London, United Kingdom*

Scott Chaplowe will present that while there is a value to standardized, quantitative approaches for impact and attribution analysis, this presentation will highlight the limitations of such approaches for the measurement and evaluation of resilience-strengthening interventions, and will highlight the importance of context-specific, participatory methods that allow resilience to be defined and measured according to community realities and needs.

**PN 62 Participatory Approaches to Resilience Measurement: Who are we Listening to?**

## **PC 233 - Participatory Approaches to Resilience Measurement: Who are we Listening to? Do we Consider Gender and Other Inequalities?**

S. Sridharan<sup>1</sup>

<sup>1</sup> *St. Michaels Hospital, Director of the Evaluation Centre for Complex Health Interventions at Li Ka Shing Knowledge Institute, Toronto, Canada*

Sanjeev Sridharan reminds us that a focus on gender and inequalities may need to consider intersections between gender, class, race, education, caste, etc. and the power relations that drive such hierarchies. How can participatory M&E and analysis illuminate such intersections and hierarchies? What can we learn by listening?

**PN 62 Participatory Approaches to Resilience Measurement: Who are we Listening to?****PC 234 - Participatory Approaches to Resilience Measurement:  
Who are we Listening to? Whose Participation and How  
to Address Conflict of Interest?**

I. Jean<sup>1</sup>

<sup>1</sup> CDA, Co-Director of Collaborative Learning and Director of Evaluation and Learning, Cambridge, USA

Isabella Jean, CDA's Listening Project reached 6,000 people in 20 countries and listened to people experiencing chronic vulnerability, slow-onset and rapid-onset disasters and conflicts, noting striking consistency in their descriptions of how aid organizations do not sufficiently understand or consider the system of relationships they are stepping into which is the cornerstone of the Do No Harm approach. Recent commitments to test flexible funding mechanisms also require an equal attention to inclusive and broad community participation in shaping priorities and engaging people in decision-making affecting their lives.

**PN 62 Participatory Approaches to Resilience Measurement: Who are we Listening to?****PC 235 - Panel Contribution**

C. McQuistan<sup>1</sup>

<sup>1</sup> Practical Action, Senior Advisor on Climate Change and Disaster Risk Reduction, Warwickshire, United Kingdom

Colin McQuistan, Practical Action UK piloted a participatory resilience measurement tool in over 100 communities in 11 countries over the last 3 years. The results from the measurements vary hence very different interventions have been planned and implemented. The differences in approaches expose different depths of engagement and early indications are that the best outcomes come from those engagements that have reached the weakest links across social, cultural, economic and environmental dimensions. The presentation will highlight some of these human stories that demonstrate very different perceptions of resilience.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 63 Applying Theory To Policy + Program Evaluation in Different Contexts: An Integrated Approach**

## Applying Theory To Policy + Program Evaluation in Different Contexts: An Integrated Approach

*S. Montague*<sup>1</sup>

<sup>1</sup> Performance Management Network Inc., President, Ottawa- Ontario, Canada

Evaluations, whether done in the development context or focused upon domestic sectors and issues has had a tendency to focus on programs and occasionally on policies. They have rarely looked at both as an integrated whole. In this session you will hear from evaluation practitioners and commissioners who will describe how they have addressed:

- 1) creating + evolving a typology of archetypical policy + program interventions;
- 2) looking at implementation or action theories as well as theories of change as seen for the primary recipients + intermediaries;
- 3) How theories of action and change can work in conditions of adaptive management.

The objective is that participants will leave this session with:

- An increased understanding of policies and programs and their related archetypical theories of change (program theories)
- A useful and practical approach to showing how policies and programs combine to create results (both intended + unintended)
- Means to incorporate reach in depictions of theories of change and in analysis and reporting. This represents a major assist to the review of gender related aspects of policy and program performance– as well as the specific analysis of various other community characteristics.
- Exposure to a generative approach to establishing, refining, collecting information on, analyzing + reporting contextualized performance – and how adaptive management can handle theories of change
- A demonstration of how integrated policy and program thinking can and should represent the essence of modern results based management principles (ref 2017 UN RBM review)
- Tools and take-aways to help people to conduct practical applications using domestic and international development contexts

Panelists will present perspectives, models and tools which should be of interest to evaluation practitioners and managers alike (Both groups will be represented on the panel.). The approach systematically applies theory around policy and program archetypes well beyond conventional practice and suggests that some archetypical configurations can be developed for various areas that will serve as touchstones for evaluation and review and, going further may help organizations to generatively learn what works for whom in what contexts and why. The approach inherently suggests particularizing before generalizing and therefore is directly relevant for gender based analysis as well as the analysis of various intersectionalities. Presenters will discuss how such an approach can be used in different contexts – including broad strategic evaluations, program evaluations and as a support to adaptive management practices. At the end of the day – the approach being demonstrated by various panelists using various application areas and contexts provides a means to understand systems interactions and to tackle issues and problems relating to resilience and sustainability.

**PN 63 Applying Theory To Policy + Program Evaluation in Different Contexts: An Integrated Approach****PC 236 - Applying Theory To Policy + Program Evaluation in Different Contexts: An Integrated Approach**

*S. Montague*<sup>1</sup>

<sup>1</sup> *Performance Management Network Inc., President, Ottawa- Ontario, Canada*

Evaluations, whether done in the development context or focused upon domestic sectors and issues has had a tendency to focus on programs and occasionally on policies. They have rarely looked at both as an integrated whole. In this session you will hear from evaluation practitioners and commissioners who will describe how they have addressed:

- 1) creating + evolving a typology of archetypical policy + program interventions;
- 2) looking at implementation or action theories as well as theories of change as seen for the primary recipients + intermediaries;
- 3) How theories of action and change can work in conditions of adaptive management.

The objective is that participants will leave this session with:

- An increased understanding of policies and programs and their related archetypical theories of change (program theories)
- A useful and practical approach to showing how policies and programs combine to create results (both intended + unintended)
- Means to incorporate reach in depictions of theories of change and in analysis and reporting. This represents a major assist to the review of gender related aspects of policy and program performance– as well as the specific analysis of various other community characteristics.
- Exposure to a generative approach to establishing, refining, collecting information on, analyzing + reporting contextualized performance – and how adaptive management can handle theories of change
- A demonstration of how integrated policy and program thinking can and should represent the essence of modern results based management principles (ref 2017 UN RBM review)
- Tools and take-aways to help people to conduct practical applications using domestic and international development contexts

Panelists will present perspectives, models and tools which should be of interest to evaluation practitioners and managers alike (Both groups will be represented on the panel.). The approach systematically applies theory around policy and program archetypes well beyond conventional practice and suggests that some archetypical configurations can be developed for various areas that will serve as touchstones for evaluation and review and, going further may help organizations to generatively learn what works for whom in what contexts and why. The approach inherently suggests particularizing before generalizing and therefore is directly relevant for gender based analysis as well as the analysis of various intersectionalities. Presenters will discuss how such an approach can be used in different contexts – including broad strategic evaluations, program evaluations and as a support to adaptive management practices. At the end of the day – the approach being demonstrated by various panelists using various application areas and contexts provides a means to understand systems interactions and to tackle issues and problems relating to resilience and sustainability.

**PN 63 Applying Theory To Policy + Program Evaluation in Different Contexts: An Integrated Approach****PC 237 - Policy and Programme Integration in Natural Resources Focusing on the Energy Sector (Including Innovation)**

M.K. LaMarche<sup>1</sup>

<sup>1</sup> *Natural Resources Canada, Evaluation, Ottawa Ontario, Canada*

This presentation will show how an internal government evaluator has evolved thinking about the evaluation policy, both in terms of strategic policy capacity (i.e., to deliver on the priorities of the government/department), and the consideration of policy in evaluations of programmes or initiatives. This includes incorporating risks, contexts, and assumptions related to program policy in the theory of change for the program. In doing so, evaluations can consider the extent to which programmes: (1) are supported by policy; and (2) align with the policy framework (i.e., in terms of delivery and intended results). From an organizational perspective, this can potentially facilitate a determination of whether the capacity exists to deliver on the policy framework, in accordance with the priorities of the government/department or if there are any gaps in policy or programme areas that would impact relevance or effectiveness. Similarly, for strategic policy areas, articulating the theory of change helps make explicit the value proposition – whether as a professional service, knowledge development, or convenor/facilitator – that can then facilitate an evaluation of policy outcomes rather than just outputs.

**PN 63 Applying Theory To Policy + Program Evaluation in Different Contexts: An Integrated Approach****PC 238 - How Theory of Change Can Work to Link Programme to Policy in Conditions of Adaptive Management**

B. Dillon<sup>1</sup>

<sup>1</sup> *Department for International Development, Research and Evaluation Division, London, United Kingdom*

This input will consider two related areas in adaptive management settings – (1) the versatility of theory of change and how it lends itself to linking programme to policy levels. (2) What to include on theory of change (ToC) when commissioning evaluations which seek to bring programme and policy closer together. The rationale for this input is that although theories of change (ToC) are core to both design and evaluation of interventions, application of ToCs has variable understanding. The objective of the session is to give the audience key insights to ToC application in adaptive management, and in particular to show how ToC lends itself to integrating or relating programme to policy levels. The session will also consider a commissioner's perspective, in evaluations which seek to bring programme and policy closer together. This topic is relevant to the wider evaluation community, particularly practitioners, managers and evaluation commissioners, who are grappling with theory of change and adaptive management. It draws on DFID experience, and the experience of the joint DFID-USAID programme on adaptive management. The topic addresses gender as the theory of change tool includes analysis of 'reach' ie whom it affects. The topic represents a relatively recent and innovative area in evaluation, and will contribute to EES participants' understanding and learning of a particular type of application, and how programme and policy levels can be usefully linked, and hence contribute to building evaluation capability. The topic is relevant to the main theme of the conference – resilience – as evaluation methodology must evolve in the face of changing demands and the increased range of data collection tools available. This type of thinking, and the combination of theory of change and adaptive management can be very powerful. It is important to understand how it works.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PN 64 From Niche to Mainstream? Critical Reflection on the Sensemaker Practice to Date

## From Niche to Mainstream? Critical Reflection on the Sensemaker Practice to Date

S. Deprez<sup>1</sup>, A. Hanchar<sup>2</sup>, I. Guijt<sup>3</sup><sup>1</sup> *Voices That Count, Training & Consultancy, Leuven, Belgium*<sup>2</sup> *The Data Atelier, Consultant, London, United Kingdom*<sup>3</sup> *Oxfam GB, Head of Research and Publishing, Oxford, United Kingdom*

SenseMaker is a narrative-based monitoring, evaluation and decision-making methodology. It is designed to help generate actionable insights and guide interventions in complex systems and processes. SenseMaker recognizes that enabling respondents to analyze and give meaning to their own experiences generates better contextualized knowledge and more reliable data. The methodology goes beyond the usual qualitative approaches, which rely on researcher interpretation. Instead, it invites respondents to self-interpret their anecdotes, experiences or stories through a set of follow-up questions at the point of origin. The responses to these questions can be used for visualisation of patterns and for quantitative analysis backed with explanatory narratives. SenseMaker can be used as a stand alone method or in combination with other assessment, monitoring, evaluation or research approaches. It is particularly useful for programmes that work in complex environments and need to understand emergent patterns and trends in real time. It provides a practical way to capture the perspectives of people and communities especially from those that are often not heard. SenseMaker is often used to understand the hard-to-measure issues such as sustainability, gender, inclusion and resilience. It generates insights in the dynamics, relationships, shocks and emerging trends in complex systems. It seeks to understand emergent properties – interactions and interdependencies that happen in an unplanned way. It is therefore an effective approach to understand resilience of people, communities or systems and allows initiatives to adapt to continuous disturbances and changing conditions. Over the last years SenseMaker has been used in a variety of international development programmes and sectors. There have been success experiences but practitioners also faced challenges during the design and implementation related to suitability to context, feasibility to implement, capacity-related issues and use findings for action. In this session, three experienced SenseMaker practitioners will critically reflect on the SenseMaker method and its use for planning, monitoring and evaluation of development programmes. By doing so we aim to contribute to an enhanced understanding of this method and ensure that when used, it adds value and achieves programme objectives. The speakers will focus on the three main questions: (1) What is the added value of SenseMaker? (2) What do we learn from practice to date? (3) From niche to mainstream? After a brief introduction to the SenseMaker method, we will move to the three panels covering each of the three main questions. Each panel will consist of a critical reflection built on the filed experience of the panellists and their colleagues, followed by an Q&A session.

**PN 64 From Niche to Mainstream? Critical Reflection on the Sensemaker Practice to Date****PC 239 - What Is the Added Value of Sensemaker?**

*I. Guijt*<sup>1</sup>

<sup>1</sup> *Oxfam GB, Head of Research and Publishing, Oxford, United Kingdom*

In this session we critically reflect on what SenseMaker contributes to the M&E practice and thinking; how it compares to other methods and whether it offers new and complementing ideas and practices for M&E processes in development programmes. Starting from the underlying principles and conceptual design of the method, we will highlight what it adds (and what not) to the way we understand reality, measure change, generate insights and inform interventions. Is it indeed relevant and can it improve our M&E practice and if yes, in what way? We discuss the typical features of the design approach, the narrative collection, the data it generates, the analysis process and use of findings and how it can complement and improve existing M&E processes. We close the session with an overview of what cannot be expected from a Sensemaker-based M&E process.

**PN 64 From Niche to Mainstream? Critical Reflection on the Sensemaker Practice to Date****PC 240 - What Do We Learn from Practice to Date?**

*A. Hanchar*<sup>1</sup>

<sup>1</sup> *The Data Atelier, Consultant, London, United Kingdom*

This session will provide a summary of the main lessons learned from the SenseMaker practice to date. Around 2010, SenseMaker was introduced in the international development sector. Initial pilots took place in programmes of organisations such as Global Giving, IRC, VECO (Rikolto), Girl Effect, and UNDP. Over the years, the interest in SenseMaker in the development sector increased with over 30 applications to date for various purposes (context studies, assessments, monitoring, evaluations and impact assessments). Based on this wide range of experiences, we will critically reflect on the practice of SenseMaker. We will present the strengths and weaknesses of the method in its use different M&E purposes and how it has been used as a stand-alone method or in combination with other methods. We will critically reflect on the practicalities of the design, collection, analysis and use, and highlight successful practice and common challenges faced by SenseMaker practitioners and in-house teams.

**PN 64 From Niche to Mainstream? Critical Reflection on the Sensemaker Practice to Date****PC 241 - From Niche to Mainstream?**

*S. Deprez*<sup>1</sup>

<sup>1</sup> *Voices That Count, Training & Consultancy, Leuven, Belgium*

While SenseMaker is often perceived as a resource intensive niche method with a steep learning curve, there is an increasing interest and awareness of this complexity-aware method. In this session, we will explore adjustments and changes required for programmes and organisations to effectively apply SenseMaker. Is there potential for SenseMaker to become a mainstream M&E method? We present our reflections, ideas and directions. We also discuss potential novel uses of SenseMaker for development and social impact initiatives.

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 65 Evaluation in Resettlement: Accountability, Vulnerability and Contested Contexts****Evaluation in Resettlement: Accountability, Vulnerability and Contested Contexts***I.L. Aronsson<sup>1</sup>, V. Gandhi<sup>2</sup>, H. Hassnain<sup>3</sup>, S. Tamondong<sup>4</sup>*<sup>1</sup> Uppsala University, Alm, Uppsala, Sweden<sup>2</sup> The Development Café, The Development Café, Jakarta, Indonesia<sup>3</sup> Founder of Pakistan Evaluation Association and Impact, Results and Learning Manager at Y Care International, Islamabad, Pakistan<sup>4</sup> UN Women, New York-as Adviser Global Evaluation Advisory Council-, Manila, Philippines

This panel discusses challenges and opportunities in monitoring and evaluation (M&E) of involuntary resettlement to improve practice, identify contexts, elements and structures that are at the core of the subject matter. The number of people exposed to DFDR (development-forced-displacement and resettlement) fluctuates between the conservatively estimated 15 million a year for reservoir driven development projects, to 250 million people in the last 20 years displaced by development projects. Climate change and conflict in fragile, violently divided societies accelerate this global conundrum. Apart from the challenge of volume, this is about human suffering, political instability, and sustainable societies. Hence, resettlement evaluation, design and operationalization, need to be refined to allow for an assessment of project implementation, decision-making and feedback on an array of issues. We need to reconsider the conventional and standard approaches to M&E in resettlement projects to include more subtle and intangible impacts. The traditional qualitative and quantitative data have to be better integrated, and we have to make use of ICT advancement to reach robust and holistic program evaluation standards emphasizing long-term outputs. This panel explores avenues for M&E in resettlement, based on the authors' long-term experiences in resettlement from the field, evaluation and research. Case studies from the Philippines, Indonesia, Mexico, Afghanistan, Jordan, DRC, South Sudan and Iraq will illustrate the complexity and urge of the problem.

**PN 65 Evaluation in Resettlement: Accountability, Vulnerability and Contested Contexts****PC 242 - Topic of Discussion: Dynamics in Resettlement and Livelihood Options for Refugees in Transition***V. Gandhi<sup>1</sup>*<sup>1</sup> The Development Café, The Development Café, Jakarta, Indonesia

Although Indonesia has not ratified the 1951 Refugee Convention on Refugees, the country has a long tradition of hosting refugees and people in need of international protection. Today, there are some 13,800 refugees registered with the UNHCR office in Indonesia. Finding an appropriate longer-term solution for each refugee is a complex and often lengthy process that involves considering the circumstances of the individual or family and identifying solutions that match their particular needs. The discussion this panel will draw from a Project of the think tank The Development CAFÉ, in Indonesia called "I AM" which is a refugee skill building project, aimed to assist refugees in transition through Livelihood skills training, particularly relating to IT Jobs as well as emotional resilience building. The discussion will also highlight challenges encountered in carrying out this project due to practical and political situations surrounding the issue of refugee resettlement.

**PN 65 Evaluation in Resettlement: Accountability, Vulnerability and Contested Contexts****PC 288 - Use of Technology in Evaluations with Displaced Population in Fragility, Conflict and Violence**

H. Hassnain<sup>1</sup>

<sup>1</sup> Founder of Pakistan Evaluation Association and Impact, Results and Learning Manager at Y Care International, Islamabad, Pakistan

This presentation will discuss the role of ICTs in achieving SDGs in countries affected by conflict and fragility. In areas where access is restricted because of conflict or warring parties, ICTs can offer creative solutions to mitigate the need for face-to-face interaction. ICTs, if combined with human interactions can play vital role in insecure environments for effective and real time monitoring of programs, evaluations of development and humanitarian actions and to acquire and respond to the feedback and complaints from the affected population. Some of the most critical factors in achieving the SDGs in countries affected by conflict are by efficiently understanding conflict dynamics to minimize harm and establishing effective partnerships to gather and analyze data with state, civil society organizations, academia and the private sector. One key reason for a lapse in this is the dearth of effective monitoring and evaluation practices and trust between different actors operating in volatile contexts due to a variety of reasons. The use of information technology has the potential to bring these institutions together by establishing and sharing responsible data that can be then used for quick decision making for a faster response in situations of emergencies. The MDGs results report show that countries that lacked behind in achieving the millennium goals were mainly the fragile states. Although there are many challenges associated to achieving results in these countries but we need to also understand that we also don't have right tools and resources available to measure progress in rapidly changing and fluid contexts. ICTs could play a key role in opening up new data approaches, for example by providing safe and economical data analytics of huge data sets; real-time feedback and greater voice for poor and marginalized groups; reduced time and cost of data collection; better data visualization for improved decision-making; and sharpened capacity for analysis of causal relations in complex scenarios. Though ICTs bring huge potential, they also carry a number of new risks, especially in conflict and fragility. This presentation will shed some light on the big-picture ethics of introducing ICTs in volatile contexts with a focus on minimizing harm and maximizing benefits to the people that are already affected by conflict. Case studies from Afghanistan, Jordan, DRC, South Sudan and Iraq will be shared to illustrate how ICTs are already being incorporated into the evaluation work.

**PN 65 Evaluation in Resettlement: Accountability, Vulnerability and Contested Contexts****PC 243 - The Challenges of Monitoring and Evaluation in Resettlement in a Calamity Affected Road Project when Design is Flawed**

S. Tamondong<sup>1</sup>

<sup>1</sup> UN Women, New York-as Adviser Global Evaluation Advisory Council-, Manila, Philippines

This is a case study demonstrating the challenges of doing M&E for a road rehabilitation project in the island of Western Samar, Philippines devastated by super typhoon Haiyan (Yolanda) in 2014. The calamity that hit the island was the deadliest typhoon ever recorded in the country and displaced thousands of people. Infrastructure needed rehabilitation and reconstruction in the island. The Millennium Challenge Corporation (MCC) provided funding to the Government of the Philippines in the form of compact agreement to rehabilitate a main artery of road network and financed an external monitoring and evaluation of resettlement among the road-affected population. The evaluation Team monitored resettlement of about 20,000 displaced people living along the road network of about 300 km. The Team conducted surveys and inter-

views in remote areas and conducted participatory consultations among stakeholders. Challenges encountered by the Team in the interview processes, sampling, survey and reporting will be discussed, including the politics of dealing with contractors, implementing agency and funder. How can evaluators deal with an evaluation, which is wrongly designed? The discussion will draw lessons learned in the process of conducting the external monitoring and evaluation in this case study.

#### **PN 65 Evaluation in Resettlement: Accountability, Vulnerability and Contested Contexts**

### **PC 244 - Evaluating the Negotiation Space in Resettlement**

*I.L. Aronsson<sup>1</sup>*

<sup>1</sup> *Uppsala University, Alm, Uppsala, Sweden*

If informed local participation is practiced, involuntary resettlement always involves negotiations between the main stakeholders – the implementor(s) and the local people. These negotiations constitute the foundation, and define the frames for restitution and compensation. In other words, they determine the future of the affected people. Research and practice have shown over time that resettled societies have difficulties to become sustainable in the future. Monitoring and evaluation (M&E) in resettlement must therefore be equipped to identify, entangle and use this highly dynamic space of negotiations that the main stakeholders jointly develop during the implementation. This space consists of power relationships, politics, and social relations, tangible and intangible dimensions of heritage, memory and culture. They are often expressed in “soft” socio-cultural categories and are difficult to discern, but are essential for learning and accountability. Too often these categories, have to be (re)formulated to fit standard approaches to program evaluation that tend to emphasize short-term outputs of activities and resource efficiency accountability, instead of a longitudinal understanding of the impacts of resettlement programs. Innovative thinking is needed to avoid repetitive negative results. The discussion will be illustrated with the ethnographic longitudinal case study of the Zimapán resettlement project in Mexico.

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 66 National Evaluation Policies and Systems in Europe: accomplishments and challenges for collective action**

## National Evaluation Policies and Systems in Europe: Accomplishments and Challenges for Collective Action

*A. Kalugampitiya<sup>1</sup>*

<sup>1</sup> EvalPartners Executive Coordinator, Executive Coordinator, Sri Lanka, Sri Lanka

A National Evaluation Policy (NEP) has a coordinating function as regards the practice of evaluation across government departments and agencies (EvalPartners, 2015). There is a great variety of NEPs depending upon the format (legislated, informal etc). National Evaluation Policies have evolved depending on country context – traditions, through various requirements of EU co-financed policies or other demands from international donors.

In this respect, the proposed panel discussion has two main objectives: 1. To discuss the challenges for the National Evaluation Policies (e.g. in terms of adaptability), with the aim of promoting resilience and action in critical times, and 2. To review the possible impact of the Global Evaluation Agenda to the formulation/reshaping of NEPs across Europe and the role of VOPEs in this process. This panel can also serve as an opportunity to establish a link with previous EES2014 discussion on NEPs, draw lessons from European states and national VOPEs that have institutionalized evaluation and developed various forms of its governance and examine the extent to which such an approach is still considered appropriate under the light of multiple crises faced in the EU. The participating experts could discuss if/how can an NEP contribute to more resilient societies and what could be the attributes of an NEP in order to respond to challenges posed by multiple financial, geopolitical, environmental and humanitarian crises. This topic is of particular relevance to the members of the evaluation community, especially to VOPEs advocating for the GEA and public sector officials involved in the shaping of NEPs and it is expected to generate a debate on the adequacy of existing NEPs to address issues of resilience in society and emerging crises, building on the relevant debate at international (e.g. EvalPartners-IOCE) and European level (e.g. 2014 EES Conference).

**PN 66 National Evaluation Policies and Systems in Europe: accomplishments and challenges for collective action**

## PC 245 - National Evaluation Policies in Europe: Observations and Learning from the Voluntary Nations Reviews (VNRs) (Kassem)

*K. El Saddik<sup>1</sup>*

<sup>1</sup> EvalSDGs Co-Chair, Expert - Policy Analysis & Evaluation, Lebanon, Lebanon

The contribution builds on the findings and observations of the latest two rounds of Voluntary Nations Reviews (VNRs) submitted to the High Level Political Forums by the European countries. VNRs intend to provide a stocktaking of the institutional, policy and implementation of the Agenda 2030 objectives. The paper will extend the work done by EvalSDGs-IIED and explore how the European countries have approached the follow up and review process at the national level. It will examine the various institutional enablers (focusing mostly on the national evaluation policy, systems and structures) that are gearing the national sustainable development agendas and map them using the Maturity tools suggested by the EvalSDGs Spotlight initiative.

**PN 66 National Evaluation Policies and Systems in Europe: accomplishments and challenges for collective action**

**PC 246 - Prospects for Developing a National Evaluation Policy in Greece: the Role of the Hellenic Evaluation Society**

L. Kantsos<sup>1</sup>

<sup>1</sup> Hellenic Evaluation Society, Board vice-president, Athens, Greece

This contribution shall refer to the case of Greece, building on the discussions of the 2017 national evaluation conference of the Hellenic Evaluation Society on how evaluation can contribute to the efforts for economic recovery and growth in the country. It shall review the responses of the Hellenic evaluation community to challenges of the crisis and examine how a National Evaluation Policy could contribute to the promotion of resilience and action in critical times. It shall also review the impact of the Global Evaluation Agenda to the efforts and options for formulating a NEP in Greece and the role of the national evaluation society in this process.

**PN 66 National Evaluation Policies and Systems in Europe: accomplishments and challenges for collective action**

**PC 287 - Towards Evidence-Based Policy Making – Institutionalization of Evaluation in Finland**

P. Uusikylä<sup>1,2</sup>

<sup>1</sup> Finnish Evaluation Society, President

<sup>2</sup> Finnish Section of the International Institute of Administrative Sciences (IIAS), President, Helsinki, Finland

Friday, 5 October 2018  
11:15–12:45

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 67 Navigating Spectacular Ambitions: The Importance of Thinking Evaluatively in Challenging the Sustainable Development Goals of “No One Left Behind”**

## Navigating Spectacular Ambitions: The Importance of Thinking Evaluatively in Challenging the Sustainable Development Goals of “No One Left Behind”

*S. Sridharan*<sup>1</sup>, *Z. Ofir*<sup>2</sup>, *J. Cekan*<sup>3</sup><sup>1</sup> *University of Toronto and St. Michaels Hospital, Health Policy- Management and Evaluation, Toronto-Ontario, Canada*<sup>2</sup> *Independent Consultant, Independent Consultant, Geneva, Switzerland*<sup>3</sup> *Valuing Voices, Valuing Voices, Prague, Czechia*

The Sustainable Development Goals (SDGs) came into effect on January 1<sup>st</sup>, 2016, with the spectacular ambition of transforming the world by 2030: “This Agenda is a plan of action for people, planet and prosperity. It also seeks to strengthen universal peace in larger freedom ... As we embark on this collective journey, we pledge that no one will be left behind.” This panel explores a simple motivating question: What will it take for the SDGs to be serious about leaving no one behind? This question implies that while it’s noble for a collective journey to leave no one behind, in a real sense there needs to be realism about the context and support structures needed to achieve such a vision. While the cold gaze of measurement and design should not dampen the enthusiasm that is inherent in a vision of leaving no one behind, we think that this panel will fill a need to think evaluatively around the types of systems, partnership networks, country-level contexts and trajectories that might influence a vision of leaving no one behind. This panel is premised on the assumption that it is in the realm of evaluative thinking to be explicit about the contexts under which a vision of no one left behind are both achievable and unachievable.

Many of the SDGs are focused on addressing inequities, yet for the most part, there is little discussion around how such inequities will be addressed across different contexts. The focus on inequities is exciting but the excitement needs to be tempered by two insights: 1) The pathways by which the SDGs propose to reduce inequities are not clear; 2) Given the wide variety of inequity gaps in different parts of the globe, there is little discussion in the SDGs on the potential heterogeneous pathways by which such inequities need to be reduced.

This panel will cover three specific areas related to the motivating question. First we ask: How can evaluations help to bring clarity to the structures, systems and partnerships across different contexts that need to be in place to achieve the vision of no one left behind? A second presentation will incorporate a view of the developmental trajectories of different countries that are experiencing very different dynamic contexts of growth and prosperity. This presentation will explore under what contexts does taking a view of leaving no one behind make realistic sense. A third paper takes a sustainability lens to the SDGs and the specific challenge of leaving no one behind. This presentation explores what would be needed for systems of solutions to be sustainable over time.

This panel is related to the theme of the conference by connecting a focus on resilience to concepts of extreme poverty and focus on individuals who live at the intersections of multiple disadvantages. It will raise questions around resilience and the ability and boundaries of sustainable solutions to address the needs of individuals who are often left behind

**PN 67 Navigating Spectacular Ambitions: The Importance of Thinking Evaluatively in Challenging the Sustainable Development Goals of “No One Left Behind”**

## **PC 247 - Incorporating Developmental Trajectories into Sustainable Development Goals**

S. Sridharan<sup>1</sup>, Z. Ofir<sup>2</sup>, J. Cekan<sup>3</sup>

<sup>1</sup> University of Toronto and St. Michaels Hospital, Health Policy- Management and Evaluation, Toronto-Ontario, Canada

<sup>2</sup> Independent Consultant, Independent Consultant, Geneva, Switzerland

<sup>3</sup> Valuing Voices, Valuing Voices, Prague, Czechia

This presentation will explore how do we think about and evaluate “leaving no-one behind” within the trajectory of development of countries in different contexts, where development models and values and principles will differ and yet this notion has to be part of “transformation” or “resilience.” The presentation will also discuss the topic from the angle of Michael Quinn Patton’s focus on principles (i.e., Principles-focused Evaluation).

**PN 67 Navigating Spectacular Ambitions: The Importance of Thinking Evaluatively in Challenging the Sustainable Development Goals of “No One Left Behind”**

## **PC 248 - Questioning the Rhetoric of “Leaving No one Behind” from a Sustainability Lens**

S. Sridharan<sup>1</sup>, Z. Ofir<sup>2</sup>, J. Cekan/ova<sup>3</sup>

<sup>1</sup> University of Toronto and St. Michaels Hospital, Health Policy- Management and Evaluation, Toronto-Ontario, Canada

<sup>2</sup> Independent Consultant, Independent Consultant, Geneva, Switzerland

<sup>3</sup> Valuing Voices, Valuing Voices, Prague, Czechia

The focus of this presentation will be on longevity. The rhetoric of “leaving no one behind” is an impossible task. International development projects target a subset of a population, and from that subset, local elites often spearhead relatively short-term projects. Jindra raises questions around what is needed both in terms of the temporalities of interventions as well as its targeted focus in order to further the aims of the SDGs – for example, broadening youth leadership, as youth potentially would keep sustaining projects long after close out. Jindra highlights examples of ex-post sustainability evaluations in Niger, Ethiopia, Zimbabwe and elsewhere.

**PN 67 Navigating Spectacular Ambitions: The Importance of Thinking Evaluatively in Challenging the Sustainable Development Goals of “No One Left Behind”**

## **PC 249 - Some Evaluation Questions to Help Sharpen the Focus of Sustainable Development Goals in Addressing Inequities: The Different Roles for Evaluators**

*S. Sridharan*<sup>1</sup>, *Z. Ofir*<sup>2</sup>, *J. Cekan*<sup>3</sup>

<sup>1</sup> *University of Toronto and St. Michaels Hospital, Health Policy- Management and Evaluation, Toronto-Ontario, Canada*

<sup>2</sup> *Independent Consultant, Independent Consultant, Geneva, Switzerland*

<sup>3</sup> *Valuing Voices, Valuing Voices, Prague, Czechia*

This presentation will explore the multiple roles that evaluation can play in enhancing the likelihood that the SDGs focus on “no one left behind” can be successful. An evaluation is much more than measurement, indicators, design and attribution. Good evaluations don't merely ask if interventions work; they raise questions about what needs to happen to make things work in different contexts. For example, realist evaluation draws attention to the context and mechanisms necessary for interventions to achieve impacts. Developmental evaluation asks: How can evaluators themselves promote the dynamic development of solutions? The utility of different evaluation approaches to the SDGs will be explored.

Friday, 5 October 2018  
11:15–12:45

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 68 Criteria for measuring resilience: what do we know and what are we missing?**

## Criteria for Measuring Resilience: What Do We Know and What are We Missing?

D. Wilson<sup>1</sup>

<sup>1</sup> Itad, Climate Change theme, Hove, United Kingdom

The evaluation community is increasingly questioning the criteria used to judge the merit and worth of interventions, welcoming a more intentional and creative use of criteria that adapts to context and need. This panel will explore evaluation criteria through a resilience lens to better understand their value for assessing resilience-focused interventions. As evaluation commissioners often include OECD-DAC criteria in Terms of References by default, we will first explore key considerations for assessing resilience interventions within this conventional configuration or criteria. We will then consider "outside this box", other criteria, principles and resilience measurement approaches to inform what and how resilience interventions are assessed. Throughout, a systemic perspective will colour our "lens," recognizing the complex interrelationships and dynamics that characterise resilience strengthening. Dave Wilson, Principle Consultant of the Climate Change Theme at Itad, will chair the session. Scott Chaplowe, Director of Evidence, Measurement and Evaluation at the Children's Investment Fund Foundation (CIFF) will use the DAC criteria as a springboard to explore key considerations for evaluating resilience. Colin McQuistan, Senior Advisor Disaster Risk Reduction and Climate Change at Practical Action, will present how principles of good development have been used to identify 'sources' of resilience, which can be evaluated to understand if project interventions are delivering resilience or not. Simone Verkaart, Monitoring and Evaluation Manager at the Global Resilience Partnership (GRP) will present how GRP's challenge round grantee have evaluated resilience on the ground, making connections back up to criteria for evaluation. Itad, as chair, will facilitate a discussion to explore the value of existing evaluation criteria for guiding resilience measurement and how we can draw from resilience building principles and approaches to inform more creative criteria for resilience measurement. Through this discussion we will aim to connect the gaps between these high-level approaches and evaluation criteria and what is measurable on the ground; and understand if and what additional criteria may be needed to comprehensively measure resilience.

**PN 68 Criteria for measuring resilience: what do we know and what are we missing?**

## PC 250 - Refocusing the DAC Criteria Through a Resilience Lens

S. Chaplowe<sup>1</sup>

<sup>1</sup> Children's Investment Fund Foundation, Evidence- Measurement and Evaluation Climate, London, United Kingdom

Evaluation as a field needs to remain discriminating, which includes questioning the value of the very criteria used to evaluate merit and worth. However, as momentum picks up to question the DAC criteria used to evaluate international development and humanitarian aid, these criteria remain the dominant recipe put forth in terms of reference, scopes of work and requests for proposals use to plan, recruit and guide evaluations. The intuitive configuration of the DAC criteria, (relevance, effectiveness, efficiency, impact and sustainability), lends to their popularity not only for development evaluation, but other evaluands. As such, while alternative criteria are welcome and can be adopted, this presentation will explore some key considerations for working with the DAC criteria when already pre-identified for the evaluation of resilience-strengthening interventions

**PN 68 Criteria for measuring resilience: what do we know and what are we missing?****PC 251 - Using Development Principles to Guide Resilience Measurement**

C. McQuistan<sup>1</sup>

<sup>1</sup> *Practical Action, Climate Change, Rugby, United Kingdom*

Colin McQuistan, Senior Advisor for Disaster Risk Reduction and Climate Change at Practical Action, will present how Practical Action have used principles of good development to identify 'sources' of resilience, and evaluated these to understand if project interventions are delivering resilience or not. Colin will introduce Practical Action's experience of the challenges to building resilience in poor developing communities. He will draw from rural and urban examples of this to explain how to allocate scarce resources when faced with multiple needs. He will explain how Practical Action have drawn from this evidence to develop a framework for measuring resilience, based on the Sustainable Livelihoods Framework. He will share the findings from the pilot of this tool in over 100 communities around the world, explaining how it can be used to understand resilience building within communities. Finally, he will reflect on the value of this tool for providing insight on resilience, highlighting lessons learned and challenges faced.

**PN 68 Criteria for measuring resilience: what do we know and what are we missing?****PC 252 - Resilience Measurement on the Ground: Experience From GRP's Grantees**

S. Verkaart<sup>1</sup>

<sup>1</sup> *Global Resilience Partnership, Monitoring and Evaluation, Stockholm, Sweden*

Simone Verkaart, Monitoring and Evaluation Manager at the Global Resilience Partnership (GRP) will present how GRP's challenge round grantee have evaluated resilience on the ground, making connections back up to criteria for evaluation. The GRP believes that resilience underpins the achievement of sustainable development in a rapidly changing world. The Challenges are a series of competitions hosted by the Resilience Partnership to tackle the world's most intractable problems. Through its Challenges, the Resilience Partnership surfaces bold, innovative ideas with real-world impact that may start small, but have the potential to scale up. GRP is keen to share lessons learnt from evaluating and managing these projects on the ground to inform wider resilience programming.

Friday, 5 October 2018  
11:15–12:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 69 The use of Causal Mechanisms in Evaluation: Collecting practical Experiences of applying different mechanism-based Approaches**

## The use of Causal Mechanisms in Evaluation: Collecting Practical Experiences of Applying Different Mechanism-Based Approaches

*J. Schmitt<sup>1</sup>, S. D'Errico<sup>2</sup>*<sup>1</sup> DEval - German Institute for Development Evaluation, Competence Centre for Evaluation Methodology, Bonn, Germany<sup>2</sup> Iee - International Institute for Environment and Development, Monitoring- evaluation- accountability and learning manager, London, United Kingdom

Faced with complex programs and multiple evaluation questions, evaluators often struggle to open the programmatic black box and uncover the causal mechanism(s) between the intervention and the observed changes in outcomes. Drawing from methodological innovations in the social sciences, evaluators keep searching for appropriate approaches to explore and explain the causal mechanisms leading to outcomes in different contexts. Advocates of mechanism-based approaches claim to have found just that. But is that true? And if it is, how are mechanisms currently used in different methods to develop deeper understanding about how programmes work? Causal mechanisms take center stage in current debates on evaluation methodology. The major added value from analysing causal mechanisms in evaluations is that they provide a better understanding of the way in which complex programmes are implemented. In addition, a stronger focus on the underlying mechanisms enables more credible statements on causal links between the contribution of an intervention and the observed effects. At the same time, the uncovering of basic causal mechanisms can increase the generalisability of results. In a nutshell, unpacking the underlying mechanisms enables more robust findings and credible statements about impact and effectiveness. In evaluation conferences and journals mechanisms are often reported as being the answer to the lack of explanations offered by the more traditional impact evaluation techniques. Especially when programmes are complex and when there are many intervening factors in the context they are operating, there is a great need to broaden the range of designs and methods for assessing causality in evaluation (Stern et al. 2012). The use of causal mechanisms as an explanatory tool spans across different methodological approaches from theory-based evaluation to statistical designs. The purpose of this panel is to compare different experiences on mechanisms-based evaluation and create a community practice to foster greater understanding about the role of mechanisms in assessing effectiveness and impact. The panel will do so by bringing together applications of different mechanism-based approaches to discuss how they can enhance evaluation practice. After a brief introduction to the topic, the panellists will provide insights into state-of-the-art mechanism-approaches from different angles of the design and methods spectrum. The panel includes recent applications of causal mechanisms in Process Tracing, Congruence Analysis, and Causal Mediation Analysis (see abstracts for individual contributions below). Following the presentations of practical examples, Dr Barbara Befani will act as discussant and provide a synthesising comment. Thereafter, the panel co-chairs will facilitate a 20–30 minutes panel-discussion. The individual contributions as well as the subsequent discussion should be guided by the following questions:

- How are causal mechanisms used across different evaluative practices? Is there a common understanding of what they are? If not, what are the commonalities and differences between different approaches?
- Can the use of causal mechanisms improve the explanatory strengths of evaluative approaches exploring impact and effectiveness? What can they add?
- How could methods based on mechanisms be combined to expand the explanatory power of different evaluative practices?

**PN 69 The use of Causal Mechanisms in Evaluation: Collecting practical Experiences of applying different mechanism-based Approaches****PC 253 - Using Mediation Analysis to Uncover Psychological Mechanisms of Attitude Change in a Development Volunteer Program**

*M. Bruder*<sup>1</sup>, *K. Guffler*<sup>2</sup>, *J.T. Polak*<sup>3</sup>, *L. Scheinert*<sup>4</sup>

<sup>1</sup> German Institute for Development Evaluation - DEval, Civil Society-Level Development Cooperation-Development Education, Bonn, Germany

<sup>2</sup> German Institute for Development Evaluation – DEval, Governance- Bi and Multilateral Development Cooperation, Bonn, Germany

<sup>3</sup> Austrian Development Agency – ADA, Evaluation, Vienna, Austria

<sup>4</sup> German Institute for Development Evaluation – DEval, Civil Society-Level Development Cooperation-Development Education, Bonn, Germany

One major aim of development volunteer services such as the German weltwärts programme is to change participants' attitudes towards host country nationals. However, the existing evaluation literature has two major shortcomings: First, there are virtually no rigorous impact evaluations examining the hypothesized positive effect of volunteering on attitudes. Second, if evaluation studies report such an effect, the causal psychological mechanisms underlying attitude change remain opaque. This is despite the fact that intergroup contact theory proposes causal pathways through which such attitude change may occur. In particular, the main factors proposed for causing more positive attitudes towards outgroups are increased knowledge about the other group, a better ability to put oneself in their position (perspective-taking), and more empathy towards them.

A theory-based evaluation of the weltwärts program conducted by DEval and published in 2017 had as one element of its design a quasi-experimental component supported by statistical mediation analysis. It thereby aimed to (a) provide a reliable effect estimate of the impact of a 1-year volunteer stay in a country of the Global South on outgroup attitudes (towards host country nationals) and (b) use mediation analysis to examine whether any positive effect could be explained by changes in knowledge, perspective-taking, and empathy. Therefore, albeit at the intraindividual "micro" level, the conceptualisation of a causal psychological mechanism conformed to the definition of any causal social mechanism. In particular, the proposed mechanisms of increased knowledge, perspective-taking, and empathy are conceptually situated between the intervention X (being a development volunteer) and the outcome Y (changing attitudes towards host country nationals) with the mechanisms supposedly being necessary conditions for X to have an effect on Y.

The evaluation observed a highly significant small-to-medium sized effect of development volunteering on attitudes of  $r = .19$ . Mediation analysis revealed that this effect was fully mediated by changes in knowledge, perspective-taking, and empathy, with each of these three pathways making unique contributions to the total effect. As such, the evaluation provides one of the first rigorous estimates of the effect of development volunteering on attitude change. However, thanks to its mechanisms approach, it did not stop at that. It also lends strong support to the validity of causal hypotheses derived from intergroup contact theory in the context of development volunteering. This mechanisms perspective adds value because of its relevance to programme design. It draws attention to the fact that being a development volunteer by itself does not change attitudes. Instead, programme managers need to design their programme such that it allows for increasing knowledge about host country nationals, seeing the world through their eyes and establishing emotional connections with them. The mediation analysis therefore substantially increased the explanatory strength of the rigorous impact evaluation by opening the black box between X and Y.

However, the challenge remains of how to use different methods of examining mechanisms systematically in an evaluation with myriads of possibly context-dependent causal psychological and social mechanisms at work and how to effectively communicate results to stakeholders of the evaluation.

**PN 69 The use of Causal Mechanisms in Evaluation: Collecting practical Experiences of applying different mechanism-based Approaches****PC 254 - Experimenting with Process Tracing Under Real-World Evaluation Constraints**

E. Raimondo<sup>1</sup>

<sup>1</sup> The World Bank, Independent Evaluation Group, Washington DC, USA

Process Tracing is a well-established method of causal inquiry in political science research and has recently made its appearance in evaluation circles, where it remains under development. Process tracing has the potential to improve the rigor and depth of evaluation findings when studying complex causal mechanisms that are highly context-specific. Real-world applications of the methods remain rare however, especially in international development evaluations. This presentation will take the audience through a practical case of applying Process Tracing to the evaluation of World Bank's support to social accountability and citizens' engagement. The presentation will focus on answering three interrelated questions:

- (i) What did the use of process-tracing uncover about the intervention and its effects that would have been undetectable with another method?
- (ii) What were the main challenges and caveats in applicability?
- (iii) What is the scope for embedding the study of causal mechanisms in large, mostly accountability driven evaluations?

**PN 69 The use of Causal Mechanisms in Evaluation: Collecting practical Experiences of applying different mechanism-based Approaches****PC 255 - Causal Mechanisms in Theory Based Impact Evaluation**

B. Wauters<sup>1</sup>, D. Beach<sup>2</sup>

<sup>1</sup> Ministry of labour and social inclusion, Director for innovation and impact evaluation, Brussel, Belgium

<sup>2</sup> University of Aarhus, Department of Political Science, Aarhus, Denmark

The presentation will contribute to discussion on how to use "mechanistic evidence" for within case causal inference in theory based impact evaluation approaches, by discussing how process tracing (as described in Beach and Pedersen, 2013) and congruence analysis (as described in Blatter and Haverland, 2012) can be used to reinforce the approaches of Chen (Change and Action models) as well as Pawson and Tilley (Realist Evaluation) within the framework of case study research as described by Yin's classic textbook. The emphasis will be on pro's and contra's of both approaches. In addition, the presentation will discuss some challenges involved in linking such approaches to comparative case study methods such as QCA.

The intervention that will serve as an illustration is referred to as a "Personal Development Process" (PDP) as described in an actual evaluation study. This study defines this PDP as a supportive process with as its goal to improve the labour market oriented personal development of the individual.

**PN 69 The use of Causal Mechanisms in Evaluation: Collecting practical Experiences of applying different mechanism-based Approaches****PC 256 - Using Realist Synthesis to Document Patterns of Effectiveness for Climate Adaptation Projects in Ethiopia**

*S. Barret<sup>1</sup>, S. Anderson<sup>2</sup>*

<sup>1</sup> *International Institute for Environment and Development - IIED, Natural Resource Group, London, United Kingdom*

<sup>2</sup> *International Institute for Environment and Development - IIED, Climate Change Group, London, United Kingdom*

Irish Aid in Ethiopia have been designing and implementing climate adaptation interventions for many years, particularly in Oromia, Tigray, and the Southern Nations, Nationalities, and Peoples' Region (SNNPR). Their objective is to enable the poor to adjust climate sensitive livelihoods to deal with current climate variability and future change. Programme implementers and assessors in multiple institutions have gained considerable experience over time of what works and doesn't work in relation to climate adaptation. However, the knowledge gained by implementers and assessors is fragmented, diffuse and typically poorly documented. Very little systematic documentation of macro assessments of multiple adaptation interventions has been conducted, as the imperative of evaluation is simply to show straightforward development impact. Further, systematic reviews of adaptation projects and programmes are impeded by the fact that standard metrics for adaptation and resilience are still yet to developed. To systematically bring evidence together and inform policy, key partners – World Forestry Centre (ICRAF), Self-Help Africa, Farm Africa, VITA, SOS Sahel, IIED among others – will conduct a realist synthesis to document and connect the linkages between adaptations interventions, their broad/narrow implementation contexts, the resultant mechanisms, and climate adaptation outcomes. A focus on causal mechanisms – connecting interventions to outcomes – is not common in the field of climate adaptation and is thus needed to improve evaluation approaches. In short, what types of climate adaptation interventions work, in what circumstances, for whom and how? The focus will be on key policy objectives within Irish Aid Ethiopia: a) the gendered impacts of climate resilient interventions; b) climate transformational agriculture; c) climate resilient seed systems; and d) equity in ecosystem based adaptation. Survey data, past assessments and personnel experience of implementers and project beneficiaries represent the information sources available. The aim is to demonstrate the key intervention-induced mechanisms of successful adaptation – whilst including the dynamics of context – and remove assumptions commonly used to connect interventions and outcomes. Focusing on the same climate resilient development outcomes, but showing different permutation and combinations of context and intervention types, assessors will aggregate all available information on climate adaptation interventions. The result will be a synthesis of findings on the effectiveness of different climate adaptation interventions, and an improved programme theory on which to design national climate adaptation programmes within Ethiopia. A further aspiration is to provide assessors and evaluators of climate adaptation projects/programmes a methodology to develop similar macro programme and policy evaluations that offer an account of important intervention mechanisms and changes in context.

Friday, 5 October 2018  
11:15–12:45

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### PN 70 Emerging Evaluation Communities in South East Europe (SEE)

## Emerging Evaluation Communities in South East Europe (SEE)

*V. Danilov*<sup>1</sup>

<sup>1</sup> Management Consulting Association MCA 2000, Macedonian Evaluation Network- MEN, Skopje, Macedonia the former Yugoslav Repu

**Abstract:** In recent years, national evaluation societies or informal evaluation networks have been established in all countries of South Eastern Europe (SEE), with the objective of strengthening local communities of evaluators and promoting evaluation culture, methods and practice.

Evaluation has been growing in SEE in line with the international and European practice. However, the lack of a structured approach to evaluation, national evaluation systems, and small potential in the use of evaluation has been identified in the SEE countries. Moreover, outside the programmes receiving international financial support, the vast majority of institutions do not use evaluation as a tool for understanding their own policies and strategies. Evaluation is not recognized as the key factor of development of good governance and evidence based policy making.

The lack of effective evaluation system results in confusion in use of progress indicators and causes difficulties in measuring the level of goal achievement, efficiency of utilization of funds. The role and function of the national evaluation societies, evaluation networks and the civil society's involvement in the field of evaluation is not understood well by those that are supposed to incorporate evaluation in their activities. The links between the governments and these organizations are weak, which makes the evaluation non-transparent and limits the benefits that it can and should provide to policy makers.

The objective of the proposed panel is to identify reasons of poor evaluation culture and practice in SEE countries and discuss ways of improving institutional capacities and individual evaluation capabilities in SEE countries.

The panel discussions aim at strengthening regional cooperation in evaluation and promoting networking among the evaluation communities in the SEE countries.

The panel will also encourage dialogue among SEE evaluation communities and with the EES. Hopefully, ways to adopt efficient practices in enhancing evaluation culture and promote evaluation in the region will be identified.

The panel will contribute to identification and elaboration of key factors of poor evaluation capacities in the SEE. The panellists are expected to come up with ideas and creative proposals for a better evaluation policy and practice in the SEE including:

- promotion of evaluation
- improvement of institutional capacities and individual capabilities of evaluators
- creating capacities for improving national evaluation capacities,
- defining national evaluation systems and policies,
- facilitating parliamentarians' involvement in evaluation and regional experience sharing,
- education and training institutions that would generate expertise with relevant staff employed in public sector, and competent evaluators in the civil sector
- improvement of individual expert knowledge for evaluation

Discussions will provide for recognition of the need for improving the public administration's performance by introducing result-based management. This will allow for better decisions about what should be replicated and what should be avoided and how corrective actions should be taken.

**PN 70 Emerging Evaluation Communities in South East Europe (SEE)****PC 257 - Panel Contribution**

M. Sumpor<sup>1</sup>

<sup>1</sup> *Institute of Economics- Zagreb, Croatian Evaluators Network- CEN, Zagreb, Croatia*

Marijana Sumpor will talk about experiences in establishing a national platform of evaluators and regional cooperation within the Western Balkan Evaluation Network and key issues of concern for evaluators in the region. She will talk about the current state of evaluation policy and practice in Croatia.

**PN 70 Emerging Evaluation Communities in South East Europe (SEE)****PC 258 - Panel Contribution**

O. Mrinska<sup>1</sup>

<sup>1</sup> *EBRD, Evaluation Department, London, United Kingdom*

Olga Mrinska will talk about evaluation elements from the EBRD perspective, commenting also on Western Balkans regional initiatives, which the Bank is supporting and promoting quite successfully. She will propose ideas that might create an additional opportunities for the evaluators and CSOs to be engaged.

**PN 70 Emerging Evaluation Communities in South East Europe (SEE)****PC 259 - How to Help Evaluation Communities Emerge when the Policy Stakeholders are Reluctant to Perform Evaluation?**

G. Michailidis<sup>1</sup>

<sup>1</sup> *Aristotle University of Thessaloniki, School of Spatial Planning and Development, Faculty of Engineering, Thessaloniki, Greece*

In South East Europe, there is an unfavourable environment for evaluation for, let us say, historic reasons. Policy makers were not performing evaluation if they were not obliged to, so that evaluation activity depended too much on the demand created by the ESIFs' interventions. But, when demand is so focused and limited, the effect on supply is that professionals tend to ring-fence their expertise and to develop their skills and abilities with their mind principally on the specific type of contracts to grab. Other stakeholders, in the public as well as the private sector, do not feel concerned by an operation which does not prove its *immediate* utility, especially if, for a very long period, the "value-for-money" principle looked more as a constraint than as a self-imposing imperative (ESIFs' money would anyway continue to pour). Finally and inevitably, in the academic sector, finding methodologies and creating tools did not seem as an "attractive" scientific domain; why to bother to engage in research when someone else was initially in goodwill and finally with magisterial manners providing you with tools more and more laborious? A promising sector for evaluation would of course be the day-to-day governance activity; but does it worth to "own" governance evaluation when (again) the EU provides you with tools and schemes and demands very specific quantified results and output?

In this environment, evaluation communities could emerge only if professionals, public and private policy stakeholders and academics realize (i) that diminishing resources demand realistic policy planning, (ii) that the nationally existing planning and evaluation abilities cannot be further developed unless in a wider context, that is to say, in the scale of the SEE at least and to start for. Economic crisis surely helps as for the first case; strengthening SEE-wide cooperation of VOPEs could be the means as for the second one.

**PN 70 Emerging Evaluation Communities in South East Europe (SEE)**

**PC 260 - Panel Contribution**

L. Cuna<sup>1</sup>

<sup>1</sup> *CEB Council of Europe Development Bank, Evaluation, Paris, France*

Based on recently completed evaluations in the Western Balkans, Luigi Cuna will talk about the experience of the Council of Europe Development Bank cooperating with national evaluators as well as the opportunities for and challenges of interfacing with national governments.

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 71 Whose Needs Count? Exploring different priorities and perspectives for evaluating resilience**

## Whose Needs Count? Exploring Different Priorities and Perspectives for Evaluating Resilience

*P. Silva Villanueva*<sup>1</sup>

<sup>1</sup> ResilienceMonitor, Director, Madrid, Spain

As development initiatives seek to address the realities of changing climates and increasing disaster risk there is a need to track progress and understand what works in building climate resilience, for whom, and why. The context-specific nature of climate and disaster resilience requires that evaluation approaches are flexible enough to accommodate diverse geographies, contexts, scales and voices. Yet within this complexity, whose needs are met? Whose voices heard? And what role does evaluation play in supporting (or not) the plurality of perspectives and needs?

Individuals and groups experiencing different positions of power are affected by different climate and disaster-related risks or by the same risks differently. There is a need for greater exploration of how we measure, evidence and understand the diverse and often divergent resilience capacities held by those individuals and groups of individuals, and potentially different impacts of resilience programmes.

This panel brings together diverse perspectives from often marginalised groups, focusing on issues such as gender, disability and informality, to explore different voices within the resilience dialogue. The aim of this session is to identify and understand various perspectives from (often-absent) voices, and their unique priorities, interests and needs. Panellists will discuss challenges and opportunities for integrating such diverse needs within evaluative work, drawing on experiences of both implementing and evaluating resilience-building projects. This session will challenge evaluators to consider whose needs count. The panel will reflect on how evaluations can help people to improve their lives and make societies more resilient, and the roles that evaluation can play to understand and better-integrate diverse perspectives particularly in the context of climatic shocks and stresses.

**PN 71 Whose Needs Count? Exploring different priorities and perspectives for evaluating resilience**

## PC 261 - Empowering People and Building Resilience to Risk Through the Lens of Inclusion

*V. Sword-Daniels*<sup>1</sup>, *K. Johnson*<sup>2</sup>

<sup>1</sup> Itad, Climate Change, Brighton, United Kingdom

<sup>2</sup> Stockholm Environment Institute, Asia centre, Bangkok, Thailand

Karlee will present findings from a recent project that focused on building the risk-response capacity of people with disabilities in Southeast Asia who are disproportionately affected by disasters, highlighting the gaps, challenges and opportunities for inclusion in resilience-building and evaluation. Attention will be called to the rationale for developing more inclusive resilience measurement that better reflects the needs of underrepresented groups.

**PN 71 Whose Needs Count? Exploring different priorities and perspectives for evaluating resilience**

## **PC 262 - Whose resilience? The Rationale for Accounting for Systemic Inequalities, Gender in Particular, and Their Intersections in Impact Evaluations**

V. Sword-Daniels<sup>1</sup>, A. Pretari<sup>2</sup>

<sup>1</sup> *Itad, Climate Change, Brighton, United Kingdom*

<sup>2</sup> *Oxfam Great Britain, Impact Evaluation, Oxford, United Kingdom*

Alexia will discuss different gender-sensitive approaches to measuring resilience capacities at the household and individual levels, trialled through Oxfam's quantitative impact evaluations in four different countries. She will focus on the rationale for and added-value of developing gender-sensitive and gender-differentiated approaches to measuring resilience capacities to enable different voices to be heard. Particular attention will be paid to the interactions of gender with other power dimensions, and the implications for building understanding of what matters to whom.

Friday, 5 October 2018  
11:15–12:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PN 72 How Can Evaluations of Governance Experiments Make Societal Transitions Manageable?**

## How Can Evaluations of Governance Experiments Make Societal Transitions Manageable?

*M. Hilden<sup>1</sup>, K. Lähteenmaki-Smith<sup>2</sup>, M. Crijns<sup>3</sup>, E. Saari<sup>4</sup>*

<sup>1</sup> Finnish Environment Institute, Professor, Helsinki, Finland

<sup>2</sup> Finnish Innovation Fund- Sitra, Public sector leadership development, Helsinki, Finland

<sup>3</sup> Open University, Public Management, Herleen, Netherlands

<sup>4</sup> Finnish Institute of Occupational Health, Qualitative Social Research, Helsinki, Finland

**Rationale:** The panel brings together scholar and practitioners of evaluation to debate and examine how evaluators and evaluations should respond to the proliferation of governance experiments

**Objectives sought:** The objective of the panel is to highlight opportunities and challenges of evaluating experiments and to identify meaningful ways forward.

**Brief narrative and justification:** In conditions of unpredictability and complexity there are less clear cut answers on how to develop and design policies. This is particularly true for the grand challenges of humanity such as climate change mitigation and adaptation, conservation of biodiversity and responding to technological disruptions.

In the search for solutions it is natural to turn to experimentation that allow for the testing of solutions on a limited scale and possibly in a confined setting. Examples include living labs, city level governance experiments, experimental policy pilots that test policy solution during a limited time to see what works. These experiments can be seen as related to polycentric governance, in which different entities in the overall governance setting can explore new solutions in their particular context. Polycentric governance is a concept going back to the work of the Nobel Laureate Elinor Ostrom who studied how partly or fully independent entities can develop functioning responses to difficult governance issues.

Polycentric governance can prepare the ground for innovations, but also raises significant challenges and opportunities for the design and management of evaluations. The experiments in a polycentric world only partly conform to rigorous design and therefore violate many of the criteria for standard statistical analysis of experiments. There are political, technical, ethical, legal and practical reasons for which the experiments are partly or completely uncontrolled, but this is no reason for neglecting their evaluation.

In this panel examples of governance experiments and their evaluation are presented. The starting point will be an overview of the nature of different governance experiments in a polycentric setting. The panel will provide insights that are of relevance for evaluation practitioners, managers, commissioners and users. It will be based on both more theoretical reflection and the presentation of specific cases of evaluating experiments. It will seek to generate a fruitful debate between the more rigorous 'classical' policy experimentation community and the more open ended emerging culture of experimentation, whose aims range from testing out novel governance solutions locally to promoting major societal transformations. By seriously considering the challenges of evaluating experiments, evaluators can contribute to the learning that leads to more resilient societies. The panel raises fundamental questions on the most useful roles for evaluation in conditions of unpredictability and complexity and discusses the design and management of evaluation and evaluation systems in organisations affected by turbulent times.

**PN 72 How Can Evaluations of Governance Experiments Make Societal Transitions Manageable?**

## **PC 263 - Phenomenal Evaluation Culture? Or Responding to Phenomenon-based Experimental Governance Through More Systemic and Responsive Evaluation**

*M. Hilden<sup>1</sup>, K. Lähteenmaki-Smith<sup>2</sup>, M. Crijns<sup>3</sup>, E. Saari<sup>4</sup>*

<sup>1</sup> Finnish Environment Institute, Professor, Helsinki, Finland

<sup>2</sup> Finnish Innovation Fund- Sitra, Public sector leadership development, Helsinki, Finland

<sup>3</sup> Open University, Public Management, Herleen, Netherlands

<sup>4</sup> Finnish Institute of Occupational Health, Qualitative Social Research, Helsinki, Finland

Contribution to the panel: 10–15 min, setting the scene for the evaluation of experiments.

**PN 72 How Can Evaluations of Governance Experiments Make Societal Transitions Manageable?**

## **PC 264 - The Influence of Evaluations in Polycentric Settings: a Literature Review**

*M. Hilden<sup>1</sup>, K. Lähteenmaki-Smith<sup>2</sup>, M. Crijns<sup>3</sup>, E. Saari<sup>4</sup>*

<sup>1</sup> Finnish Environment Institute, Professor, Helsinki, Finland

<sup>2</sup> Finnish Innovation Fund- Sitra, Public sector leadership development, Helsinki, Finland

<sup>3</sup> Open University, Public Management, Herleen, Netherlands

<sup>4</sup> Finnish Institute of Occupational Health, Qualitative Social Research, Helsinki, Finland

Contribution to the panel: 10–15 min, focus on the particular challenges of the polycentric setting and its effects on the potential influence of the evaluations (Also submitted as individual contribution).

**PN 72 How Can Evaluations of Governance Experiments Make Societal Transitions Manageable?**

## **PC 265 - Evaluation of Service Experiments Using Human Centered Value Criteria**

*M. Hilden<sup>1</sup>, K. Lähteenmaki-Smith<sup>2</sup>, M. Crijns<sup>3</sup>, E. Saari<sup>4</sup>*

<sup>1</sup> Finnish Environment Institute, Professor, Helsinki, Finland

<sup>2</sup> Finnish Innovation Fund- Sitra, Public sector leadership development, Helsinki, Finland

<sup>3</sup> Open University, Public Management, Herleen, Netherlands

<sup>4</sup> Finnish Institute of Occupational Health, Qualitative Social Research, Helsinki, Finland

Contribution to the panel: 10–15 min, examining the specific conditions for and opportunities in evaluating experiments for novel services. Evaluation of service experiments and how they could be evaluated by human centered value criteria and methods which make participants learn by evaluating and change their prevailing activities.

**PN 72 How Can Evaluations of Governance Experiments Make Societal Transitions Manageable?**

## **PC 266 - Frozen Experiments and the (Lack Of) Evaluability of Policy Instruments: A Case Study of Danish Work Place Assessments Under EU**

*P. Dahler-Larsen<sup>1</sup>*

<sup>1</sup> University of Copenhagen, Department of Political Science, Copenhagen, Denmark

Policy instruments have careers. They have short or long lives. Each policy instrument can in itself be seen as an experiment. But it is not always subject to experimental thinking and evaluation.

While the purpose of Evaluability Assessment (EA) is normally to distinguish between situations where evaluation is appropriate and situations where it is not, this paper uses EA as an theoretical-analytical tool to further explore the policy/evaluation interface. It does so through a case study of Danish work place assessments (“APV”), which take place with reference to EU framework directive 89/391, which has basically remained unchanged over almost 30 years. The paper thus contributes to the debate about evaluation of experiments by looking at factors which perhaps inhibit evaluative thinking and practice.

A research project on APV has been financed by the Danish Working Environment Fund. The project includes a systematic attempt to determine the evaluability of APV as a policy instrument in the context of the above-mentioned directive, using eg. the more recent evaluation of the framework directive, country reports, and other documents.

While it is shown that some technical and methodological obstacles to evaluation can be overcome by innovative evaluation manoeuvres, it is also argued that some dimensions inherent in the design of the APV legislation itself constitute fairly fundamental challenges for evaluation. Paradoxically, it is shown that while the purpose of the policy undergirding the framework directive is to produce evaluative data, these are not put to active use at the policy level, and the evaluability of the policy itself remains low. The relatively low political priority given to occupational health and safety policies may also help explain why little has been done over 30 years to increase the evaluability of the framework directive. In that sense, the framework directive 89/391 and the Danish APV legislation are frozen experiments.

#### **PN 72 How Can Evaluations of Governance Experiments Make Societal Transitions Manageable?**

### **PC 267 - Evaluating Experiments for Polycentric Climate Governance**

*M. Hilden<sup>1</sup>, K. Lähteenmaki-Smith<sup>2</sup>, M. Crijns<sup>3</sup>, E. Saari<sup>4</sup>*

<sup>1</sup> Finnish Environment Institute, Professor, Helsinki, Finland

<sup>2</sup> Finnish Innovation Fund- Sitra, Public sector leadership development, Helsinki, Finland

<sup>3</sup> Open University, Public Management, Herleen, Netherlands

<sup>4</sup> Finnish Institute of Occupational Health, Qualitative Social Research, Helsinki, Finland

Contribution to the panel: 10–15 min, raising the specific challenges related to the evaluation of the upscaling and possibilities for transformative change through experimentation and discussion of how in particular a realistic evaluation approach should be developed to deal with experiments for societal transitions. Polycentric governance has been characterized to include local action, mutual adjustment, experimentation, the building of trust and the gradual emergence of overarching rules (Jordan et al. 2018). Especially loosely defined ‘experiments’ have emerged as a way to identify and foster innovative policy actions. This has become particularly prominent in the fields of sustainable development and in actions aiming to achieve climate change mitigation and adaptation. The polycentric turn and the proliferation of experiments have raised new types of problems for evaluations. Experiments that are based detailed and experimental designs that include randomisation and rigorous controls, are fairly straightforward to evaluate, which is one reason for them gaining popularity as a base for policy development (James et al. 2017). However, they are not capable of capturing the full diversity of the needs of climate governance that is concerned with achieving societal transitions. This has led to an emerging polycentric experimental culture that includes experimentation with little or no control and piloting at different levels of governance in highly context dependent settings. They are much more difficult than the rigorous experiments from an evaluator’s perspective. Evaluators cannot, however, dismiss this activity as it is potentially important. It may lead to the justification of policies and steer policy development in many ways. Therefore it is important to explore how the evaluations of these experiments should be organized. In particular, what is the role of the evaluation in the experimentation itself and its subsequent upscaling from a local and national activity to EU wide policies? This contribution explores how a polycentric ‘culture of experimentation’ for climate governance can be approached. The analysis discusses in particular how the realistic experimentation approaches can be modified to accommodate the specific characteristics of polycentric policy experiments. as a contributor to policy innovation and policy development. The review is based on cases of experimentation in the field of sustainable

development and climate change, with particular reference to the increasingly challenging climate goals and the objectives of Agenda 2030. The aim is to critically discuss how evaluations and evaluators can contribute to the emerging culture of experimentation. James et al., 2017. *Experiments in Public Management Research*. Cambridge University Press. Jordan, A., Huitema, D., van Asselt, H. and Forster, J. 2018. *Governing Climate Change The Promise and Limits of Polycentric Governance*. In Jordan, A., Huitema, D., van Asselt, H. and Forster, J. (Eds.) *Governing Climate Change – Polycentricity in Action?* Cambridge University Press.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 73 90-minute introduction to Principles-Focused Evaluation**

## 90-minute Introduction to Principles-Focused Evaluation

*R. Wilson-Grau*<sup>1</sup>

<sup>1</sup> *Ricardo Wilson-Grau Consultoria em Gestao Empresarial Ltda, Independent Consultant, Rio de Janeiro, Brazil*

Principles inform and guide choices in the face of complexity. They do so by telling us how to act. Principles-focused evaluation examines (1) whether principles are clear, meaningful, and actionable, and if so, (2) whether they are actually being followed and, if so, (3) whether they are leading to desired results.

Principles-focused evaluation builds knowledge about which principles are appropriate for what purposes in which contexts, helping those involved in change navigate the treacherous terrain of conflicting guidance and competing advice. What principles work for what situations with what results is an evaluation question. Thus, from an evaluation perspective, principles are hypotheses not truths. They may or may not work. They may or may not be followed. They may or may not lead to desired outcomes. Whether they work, whether they are followed, and whether they yield desired outcomes are subject to evaluation

Michael Quinn Patton is an independent evaluation consultant with 45 years' experience and former president of the American Evaluation Association. He was a founding keynote speaker of AfrEA. He is the author of 6 major evaluation books including the two cited above. Principles-focused evaluation is Patton's latest contribution to the field of evaluation. In 2017, his book *Principles-Focused Evaluation for Principles-Driven Programs and People* was published. In it he explains why principles-driven programs are designed to be highly adaptive, innovative, and responsive under conditions of complexity, this new approach treats program principles as the unit of analysis (evaluand) and designs an evaluation to assess both implementation and consequences of program principles.

Principles are succinct, pointed, and specific enough to provide program direction but open enough to be adapted to context. Principles-focused evaluation supports principles-driven program design.

Then Ricardo Wilson-Grau is an international evaluator who has become deeply engaged with principles-focused developmental evaluation, contributing to both its theory and practice. He will describe how articulating principles helped define and illuminate Outcome Harvesting, an innovative evaluation mode that collects evidence of what has been achieved by an intervention and works backward to determine whether and how the efforts of social innovators and their interventions contributed to observed and documented changes. The presentation is based on a chapter Wilson-Grau wrote for Patton's *Principles-Focused Evaluation on Outcome Harvesting* and on Wilson-Grau's own book *Outcome Harvesting for Monitoring and Evaluation: – Practical Applications of Essential Principles* that will be published by IAP to coincide with the EES 2018 Conference.

These two presentations will be followed by 45 minutes of questions and answers.

**PN 73 90-minute introduction to Principles-Focused Evaluation**

## **PC 269 - Outcome Harvesting – A Principles-Focused Evaluation Approach**

*R. Wilson-Grau<sup>1</sup>*

<sup>1</sup> *Ricardo Wilson-Grau Consultoria em Gestao Empresarial Ltda, Independent Consultant, Rio de Janeiro, Brazil*

Ricardo Wilson-Grau will present in 15 minutes an example of a principle-focused evaluation approach. The two presentations will be followed by 45 minutes of questions and answers.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

#### PN 74 Visualising Theories of Change

### Visualising Theories of Change

*S. Powell*<sup>1</sup>

<sup>1</sup> *Freelance, Consultant, Clevedon, United Kingdom*

This panel is at the intersection of two trending topics in evaluation. Data visualisation: evaluators have to present their results in an accessible way, and good visuals can complement a narrative and provide new insights. Theories of Change, whether presented as visuals or text, are central to many modern understandings of evaluation and its tasks. But what happens at their intersection? How do we, should we, visualise Theories of Change? Can we apply what we have learned from visualising data to visualising theories? How do informal approaches to sketching/visualising the processes and results of participatory workshops fit into this picture? If graphical Theories of Change don't (primarily) visualise data, what do they represent? Causal networks (Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge Univ Press.)? Could visualisations of causal connections be useful to evaluators apart from presenting formal, complete Theories of Change? Is there a role for a common visual language or alphabet in the visualisation of data? What about in the visualisation of Theories of Change? Could there be a common language across both? What would be the benefits and drawbacks? How can adaptive mechanisms be represented graphically? How can we deal with interactive, dynamic Theories of Change which allow desk-based experiments that let us see the consequences of different variations in the design of theories of change? This panel welcomes both practical presentations of tips and case studies as well as more theoretical papers.

#### PN 74 Visualising Theories of Change

### PC 270 - Representing Theories of Change: A Technical Challenge with Evaluation Consequences

*R. Davies*<sup>1</sup>

<sup>1</sup> *Freelance, Consultant, Cambridge, United Kingdom*

This paper was recently commissioned by CEDIL, a DFID funded initiative expected to develop and test innovative methods for evaluation and evidence synthesis. The paper looks at the technical issues associated with the representation of Theories of Change and the implications of design choices for the evaluability of those theories. The focus is on the description of connections between events, rather the events themselves, because this is seen as a widespread design weakness. Using examples and evidence from two sources (Google Image searches and an email list discussion) six structural problems are described, along with their consequences for evaluation. The paper then outlines six different ways of addressing these problems, which could be used by programme designers and by evaluators. These include the use of specific software tools and kinds of participatory processes. The paper concludes with some caution, speculating on why the design problems are so endemic but also pointing a way forward. Three strands of work are identified that CEDIL and DFID could invest in to develop solutions identified in the paper, building on existing practice and exploring new approaches used in other fields. Together, these three strands of work, plus the advice already provided in this paper, should provide a range of solutions that address a range of representational problems, from the most basic (no identifiable connections between events in a theory of change) to the more sophisticated (no means of identifying the iterated consequences of the connections within a theory).

**PN 74 Visualising Theories of Change****PC 271 - To the Promised Land: The case of The Salvation Army on its Quest for the 'Perfect' Theory of Change**

A.M. Brown<sup>1</sup>

<sup>1</sup> Freelance, Consultant, The Hague, Netherlands

'What is 'fruit of the spirit' doing on a Theory of Change?', 'Shouldn't we have a Theology of Change instead? These are some of the questions members of a faith-based organisation grappled with as they were trying to articulate how their development work leads to change.

This article describes the journey of The Salvation Army (TSA) in developing a Theory of Change that:

- 1) balances theological epistemology with traditional development principles;
- 2) is reader ('visually') friendly to the congregation, a secular audience, technical entities such as donors and general members of the public;
- 3) facilitates effective monitoring and evaluation;
- 4) guides the design of programme and projects;
- 5) links well to other organisational tools.

Essentially a Theory of Change that is multi-purpose, practical, useable, adaptable and evaluable.

A frank description of the missteps and the lessons learnt along the way is given, as TSA went from developing a very rudimentary TOC with several design flaws such as not being evidenced-based, no outcomes pathway, no causal links, no weighting of the connections between events, viewing change as just a linear process, no feedback loops and no clear statement of the problem to adopting a TOC though far from perfect, at least satisfies in part the five conditions above.

In the article both the process for designing the TOC as well as how the content of the TOC was derived are discussed. Practical tips are also given on the use of software for TOC visualisation and group facilitation techniques to produce a TOC that both the ideological purist and the empirically-minded stakeholders within TSA can appreciate.

**PN 74 Visualising Theories of Change****PC 272 - Counting What Counts: Value in Theories of Change**

S. Powell<sup>1</sup>

<sup>1</sup> Freelance, Consultant, Clevedon, United Kingdom

This paper starts by looking at the difference between a plain "theory" and a "theory of change" (ToC). I suggest that a theory of change is a stakeholder's theory about what they have to do to get what they want. I introduce a "heart" symbol to mark ToC elements which a stakeholder values and an "action" symbol to show elements which they can control. These symbols can be used in any ToC. Making these two things explicit and visible frees us from some of the assumptions implicit in most templates and guidelines for ToCs, in particular that only variables at or near the end of a causal path can be valued/valuable (and everything else is just a means to an end). This opens up some big questions, practical, ethical and theoretical: Where do valued variables appear in theories of change? I argue that is not true that a ToC can only have one valued variable or that only variables at the end of a causal path can be valued. We also go on to look at the practical consequences in terms of project design and implementation if a template does not allow us to mark as valuable some things or variables which in fact we actually value? If there is more than one valued variable, how do we / should we give them different weights or priorities? What if some things are measured in dollars and some in "lives saved"? Can we combine cost and value in the same framework? I present a partial solution using "heart" symbols. If the valued variable stretches over time, (for example, providing the children with school meals every schoolday) how do we usually aggregate this value, e.g. when conducting an evaluation? What are the problems if we only look at the "endline"? What cognitive biases

threaten our judgement (see Kahnemann, D. (2011). Thinking, fast and slow.)? What if someone says "it isn't only about the result, it's also about the process. For example, simple acts of kindness are important in themselves."? We look at ways to value a process in the same way as a result. What about multiple stakeholders (see also Hansen, M. B., & Vedung, E. (2010). Theory-Based Stakeholder Evaluation. *American Journal of Evaluation*, 31(3), 295–313)? What if different stakeholders value different things (and what if they control different things, or even have a different theories)? – I suggest that we can use the same "action" and "heart" symbols to help combine multiple theories into one. The paper also looks at systematically different perspectives of men and women, "beneficiaries" and implementers, illustrated by a real-life example from an Ebola "Lessons learned" study which illustrates some of these problems and a partial solution. The paper concludes by reviewing the suggested "action" and "heart" symbols with some practical guidelines for how to introduce them in actual theories of change.

#### PN 74 Visualising Theories of Change

### PC 273 - More Ideas for Making Theories of Change More Informative (And Robust) Using Data Visualization

S. Vaca<sup>1</sup>

<sup>1</sup> *Freelance, Consultant, Montech France, France*

Starting with a suggested definition of what a robust Theory of Change is, and its elements, this session will be based on the soon-to-be published paper: Dhillon, L., Vaca, S. (2018). "Refining Theories of Change: Important Elements, Diverse Representations and Tool Alignment". *Journal of Multidisciplinary Evaluation (JMDE)*. Focusing on the diagrams representing them, based on a desk review of current ways of presenting ToCs, the session will highlight several strategies to make the diagrams more informative, using the same space (typically one page), but making visual pre-attentive features (such as colors, size, etc.) in a more purposeful manner that provides more information about the theory. At the same time, by making the representation more transparent, the ToC elements and their relationships will be more visible, and this could help to strengthen it. Some of the ideas that will be discussed: more meaningful causal strands, introducing an area of accountability, unintended or unexpected effects or results, Theory of Change rubric, additional elements including more information in the Causal Links, considering the level of complexity in the Theory of Change, representing visually the importance of each strategy, and making mechanisms more transparent. Finally, some mention will be made of how to align Theories of Change with other organizational tools.

Friday, 5 October 2018  
13:45–15:15

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PN 75 United Nations' evaluation functions for more resilient societies**

## United Nations' Evaluation Functions for More Resilient Societies

*M. Segone*<sup>1</sup>

<sup>1</sup> UNFPA - United Nations Population Fund, Evaluation Office, New York, USA

After a decade of turbulence in Europe and the world, we need to look toward the future and focus on the contribution of evaluation to creating more resilient societies. Crises inherently lead to change, positive and negative. What is the role of evaluation in understanding the multiple crises – economic, financial, humanitarian, social, political ... – currently ongoing? What responses does evaluation propose? The 2030 Agenda for sustainable development brings mainstream the development and humanitarian nexus, shaping the “resilience” discourse and action for the next 15 years. All countries, including developed countries in Europe as well as the UN system, are important stakeholders. The panel – composed of Directors of evaluation from 7 UN agencies – will examine the role of evaluation functions in different UN agencies in strengthening resilient societies, by addressing the following aspects:

- Challenges and opportunities for the evaluation field
- Dilemmas and trends in professionalism, standards and ethical norms
- Values in evaluation: philosophical, ethical and theoretical considerations
- Maintaining independence, relevance and responsiveness
- The role of partnerships and stakeholders
- Communicating, using and embedding evaluation

**PN 75 United Nations' evaluation functions for more resilient societies**

## **PC 274 - Panel Contribution**

M. Segone<sup>1</sup>

<sup>1</sup> UNFPA - United Nations Population Fund, Evaluation Office, New York, USA

**PN 75 United Nations' evaluation functions for more resilient societies**

## **PC 275 - Panel Contribution**

A. Cook<sup>1</sup>

<sup>1</sup> WFP, World Food Programme, Evaluation Office, Rome, Italy

**PN 75 United Nations' evaluation functions for more resilient societies**

## **PC 276 - Panel Contribution**

I. Naidoo<sup>1</sup>

<sup>1</sup> UNDP, Director, Independent Evaluation Office, New York, USA

**PN 75 United Nations' evaluation functions for more resilient societies**

## **PC 277 - Panel Contribution**

O. Garcia<sup>1</sup>

<sup>1</sup> IFAD, Director, Independent Evaluation Office, Rome, Italy

**PN 75 United Nations' evaluation functions for more resilient societies**

## **PC 278 - Panel Contribution**

M. Spilsbury<sup>1</sup>

<sup>1</sup> UNEP, Director, Evaluation Office, Nairobi, Kenya

**PN 75 United Nations' evaluation functions for more resilient societies**

## **PC 279 - Panel Contribution**

A. Ruiz Villalba<sup>1</sup>

<sup>1</sup> World Intellectual Property Organization (WIPO), Head Evaluation Section, Internal Audit and Oversight Division, Geneva, Switzerland

Friday, 5 October 2018  
13:45–15:15

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 76 Lessons Learnt In Including Young And Emerging Evaluators For A Resilient Field. A Cross Regional Exchange**

## **Lessons Learnt in Including Young and Emerging Evaluators for a Resilient Field. A Cross Regional Exchange**

*G. Sánchez Romero<sup>1</sup>, A. Bolaños<sup>2</sup>, B. Julia<sup>3</sup>, G. Jérôme<sup>4</sup>, F. Hoosen<sup>5</sup>*

<sup>1</sup> ACEVAL, Board member, Mexico City, Mexico

<sup>2</sup> EvalYouth Costa Rica, Member, San José, Costa Rica

<sup>3</sup> European Evaluation Society, TW5 - Emerging Evaluators, Prague, Czechia

<sup>4</sup> IDEAS, Board member, Luxembourg, Luxembourg

<sup>5</sup> EvalYouth Africa, VOPE initiative, Pretoria, South Africa

The Global Evaluation Agenda 2016–2020 recognizes the importance of strengthening local evaluation capacities amongst young and emerging evaluators (YEEs), in order to better implement and evaluate the sustainable development goals and local policy agendas. This call was launched in 2015, and since that date, several organizations aimed at advancing the field of evaluation around the globe undertook different initiatives to achieve these goals.

As we approach the year 2020, it is important to examine what have been the strategies that have proven to be successful at strengthening and including YEEs across different contexts globally. But more importantly, we need to find effective ways to collaborate and share lessons learnt between global regions for these strategies to be sustainable and more resilient in the near future.

For this end, we have assembled a panel that is diverse and well represented internationally. It will be composed by representatives of the EvalYouth regional chapters of Latin America and Africa, the Thematic Working Group of Emerging Evaluators from the European Evaluation Society, and from the initiatives addressed to young and emerging evaluators internationally lead by IDEAS. Thus, this panel will allow a dialogue between diverse initiatives in different regions of the world.

The goals of the panel consists on identifying success stories in strengthening evaluation capacities amongst YEEs in different regions, and exchange best practices, based on the specificities of each context. The panel is also aimed at reflecting the main lessons learnt by the implementers of these initiatives, as well as drawing the main challenges faced in designing, implementing and continuing these initiatives addressed to YEEs locally and globally.

We hope this panel will establish a continuing dialogue and define effective ways to collaborate between these regions, in order to support young evaluation professionals around the globe, and thus advancing the evaluation practice, which is stated in the Global Evaluation Agenda 2016–2020.

**PN 76 Lessons Learnt In Including Young And Emerging Evaluators For A Resilient Field. A Cross Regional Exchange****PC 291 - The Experience of Including Young and Emerging Evaluators from Evalyouth Costa Rica**

*N. Salas*<sup>1</sup>

<sup>1</sup> *EvalYouth Costa Rica, Leadership, San José, Costa Rica*

In recent years, Costa Rica has been rethinking the way in which evaluation is done, identifying the need to incorporate new approaches and experiences.

Thus, during the Solís Rivera presidential administration, a National Assessment Agenda was proposed, and for the first time, a total of eighteen strategic evaluations of National Programs and Strategies are carried out. It is at this juncture where the participation of young and / or emerging evaluators (YEEs) takes center stage in evaluation at the national level.

International Cooperation and the EvalYouth initiative has promoted the incorporation of YEEs in this Agenda, which has allowed young evaluators to put their theoretical knowledge into practice and, at the same time, to contribute in a novel and innovative ways in the methodologies used to approach evaluations. Success stories have occurred when the experienced evaluators, who lead the evaluation teams, give space to the innovation, imagination and new energy brought by the YEEs.

In this stories of success there have been experienced people with great ideas, and when they have added YEEs to their teams, the results are the elaboration of innovative and dynamic proposals. On the other hand, there is also the experience of those who consider YEEs as mere assistants in evaluation projects; limiting the capabilities of the YEEs, and making it a frustrating experience for those who wish to continue growing.

Among the good practices, it is important to consider the active involvement of YEEs in the evaluation process as an equal, considering that she is a person trained in evaluation, full of desires to contribute and learn how to transcend the theory to the practice. In evaluations where this opportunity has been granted to YEEs, the teams are strengthened, the junior and senior evaluators grow together, and a real accompaniment is provided to the junior evaluator. In addition, it transcends the "young as assistant" take towards a more proactive role.

Confidence, clear and transparent communication and the delegation of tasks allow the professional to display all her potential; an aspect that impact the team's ability to achieve results faster and achieve its objectives more effectively.

In this process of strengthening YEEs capabilities, the challenges that we still have as YEEs participating in EvalYouth Costa Rica are: encouraging more coordinators to incorporate more YEEs in their evaluation teams, and promote the message that YEEs can assume more proactive roles. Also, a barrier to overcome is the often present adult-centrism, that has limited the possibilities of young professionals who want to learn and contribute.

Thus, this paper will present the experience of five YEEs from Costa Rica who have been participating in EvalYouth Costa Rica, both from the institutional, international cooperation and evaluation teams; who will share their lessons learnt and challenges to contribute to the culture of evaluation in the country and the Latin American and the Caribbean region.

**PN 76 Lessons Learnt In Including Young And Emerging Evaluators For A Resilient Field. A Cross Regional Exchange****PC 280 - Perspective from Africa. Experiences, Successes, Challenges and Motivations for Continued Evaluation Capacity Building Amongst Young and Emerging Evaluators**

H. Fazeela<sup>1</sup>, N. Ngwabi<sup>2</sup>

<sup>1</sup> EvalYouth Africa/AfrEA YEEs Network, Task Force VOPEs, Pretoria, South Africa

<sup>2</sup> Stellenbosch University, CREST, Cape Town, South Africa

This presentation aims at sharing experiences, success stories and challenges from the EvalYouth initiative and the AfrEA YEE network in the African region, on the existing opportunities on strengthening evaluation capacities amongst young and emerging evaluators (YEEs).

Any best practices and contextual factors in enabling and hindering the creation of a YEE network will be shared. This will be at an opportune time as the AfrEA YEE network was launched in October 2017 at the SAMEA conference and is still in its initial stages of formation. The main purpose of the network is centred on strengthening evaluation capacities in the region through encouraging activities, networking, sharing of experiences, organising activities and encouraging VOPEs to integrate the issues of YEE in their strategic plans.

This exchange will highlight the role of a YEE network in facilitating the process of strengthening evaluation capacities at a regional level and what can be learnt from international and other regional initiatives. This will further inform the action plan of the AfrEA YEE network on other existing opportunities for strengthening evaluation capacities amongst YEE within the region. Experiences from the AfrEA YEE Network on the challenges faced in designing and implementing initiatives will be shared. Some successes that can be shared is fostering the need for improved capacities of YEE and the prompt application of skills obtained to nurture and strengthen the evaluation practice in Africa.

As the co-chair of EvalYouth Task force 1, I will also be in a position to share some of the current work done by EvalYouth in terms of its mentorship programme, YEE virtual conferences, and motivating YEE to become more involved in VOPE governance, committees and activities. The opportunities, challenges and successes of this work will be drawn upon to see how it has informed or can inform work at the regional level. One challenge to highlight from the mentorship programme is drop out of mentors due to job mobility and addressing these dynamics. With regard to engaging YEE in VOPEs a challenge that exists is making voluntary participation in VOPE governance to be attractive to YEE in spite of competing needs of building capacities and gaining experience in the field.

The panel discussion will therefore allow for synergies across regions to be discussed that can inform future collaborations. This will inform the approaches to strengthening evaluation capacities in Africa and across the regions that will ultimately allow YEE to be more resilient in facing complexity.

**PN 76 Lessons Learnt In Including Young And Emerging Evaluators For A Resilient Field. A Cross Regional Exchange****PC 281 - Creating Opportunities for Emerging Evaluators for a Better Evaluation Field. The Ideas Approach**

J. Gandin<sup>1</sup>

<sup>1</sup> IDEAS, Board Member, Luxembourg, Luxembourg

Emerging Evaluators are drawing a growing interest within the evaluation community. Some professional associations (AEA, CES, SAMEA, EES, EvalPartners, etc.) already integrate emerging evaluators as part of their strategic priorities and work plan and offer activities to their members, including mentoring programmes.

In 2015, IDEAS has conducted a survey whose the results claimed for giving opportunities and allow emerging evaluators to participate in projects and/or let them offer services, on a voluntary basis, for the association and its members.

Since then, IDEAS has been actively involved in partnerships with other professional associations and networks in evaluation, and fully supports existing mentoring programmes.

In a nutshell IDEAS considers networking and mentoring between young and emerging evaluators and experienced professionals are particularly crucial for stimulating and sharing knowledge and expertise amongst evaluators. This panel will present the IDEAS approach to strengthening capacities among emerging evaluators, and aims at establishing a collaboration that will enable the advancement of the evaluation practice.

#### **PN 76 Lessons Learnt In Including Young And Emerging Evaluators For A Resilient Field. A Cross Regional Exchange**

### **PC 282 - Presenting the EES Thematic Working Group for Emerging Evaluators (TWG5). The Horizon for Emerging Evaluators in Europe**

*M. Branco*<sup>1</sup>

<sup>1</sup> [www.mariana-branco.com](http://www.mariana-branco.com), Evaluation Specialist, Porto, Portugal

The EES Thematic Working Group for Emerging Evaluators (TWG 5) was created shortly after the Dublin Conference in 2014. It aims at enhancing the involvement of young and emerging evaluators in EES activities, promoting the exchange between experienced and less experienced evaluators and connecting young/new evaluation professionals and researchers across Europe and beyond. Since its establishment, TWG 5 has run a number of activities with and for emerging evaluators, including a virtual conference, a mentoring program and various capacity building initiatives. The reach of TWG 5 goes far beyond Europe and collaboration with other YEE initiatives has always been at the heart of its work. This presentation will take stock of what has been done. It will discuss what have been some of the successes and challenges. And it will ask the question of how TWG 5 may continue its work, in complementarity with other existing YEE initiatives around the globe.

Friday, 5 October 2018  
13:45–15:15

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PN 77 From Independent Evaluator To Learning Partner: Reflections On The Journey**

## **From Independent Evaluator to Learning Partner: Reflections on the Journey**

*L. Vogel*<sup>1</sup>

<sup>1</sup> *Itad, Organisational Effectiveness, Edinburgh, United Kingdom*

As funders and implementers seek to solve increasingly complex, inter-related problems many are seeking evaluators to undertake the role of a learning partner working alongside, and sometimes embedded in, portfolios, programmes and initiatives helping them to ground their strategy in sound research and then supporting them to evolve and adapt in real time based on the best available evidence. For a growing number of actors, the idea of an evaluator as an independent actor passing judgement on a programme has become anachronistic. What they want is part evaluator, part facilitator, part strategy consultant. This session draws together Itad's collective experience of serving as the learning partner to a number of clients and will engage participants in reflecting on both the opportunities and challenges of playing this role.

This panel will draw together Itad's experiences in playing the role of a learning partner with a range of clients from the public sector and philanthropic foundations. The themes to be covered will come from Governance, Organisational Effectiveness, Health, Private Sector Development and Climate Change. The evaluation methods that will be discussed include: developmental evaluation, utilisation focused evaluations and theory based evaluation.

The issues that will be explored in the session will include:

- **Supporting greater evidence use** – how being embedded within an implementation or portfolio team can help support greater uptake and use of evidence in decision making, but also the conditions that need to be in place for this to happen.
- **Being responsive while maintaining quality** – how to balance the demands of providing real time data to inform course correction, with ensuring evidence is sufficiently robust to provide the basis for decisions.
- **Navigating different roles** – as a learning partner you are often required to undertake different roles at different points in time: sometimes you are an evaluator, other times a researcher, other times advisor. Shifting between these roles requires the ability to understand culture, people and act with diplomacy and skill.

Panelists will provide 15 minute reflections on their experiences of being involved in learning partnerships. Each will provide context to the contract and reflections on the opportunities and challenges of playing the role of learning partner.

**PN 77 From Independent Evaluator To Learning Partner: Reflections On The Journey**

## **PC 283 - Panel Contribution**

*R. Lloyd*<sup>1</sup>

<sup>1</sup> *Itad, Organisational Effectiveness, Hove, United Kingdom*

**PN 77 From Independent Evaluator To Learning Partner: Reflections On The Journey**

## **PC 284 - Panel Contribution**

*S. Wallach*<sup>1</sup>

<sup>1</sup> *Itad, Organisational Effectiveness, Hove, United Kingdom*

# Round-tables

Wednesday, 3 October 2018  
10:00–11:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**RT 01 Focusing sessions: a stakeholder engagement and evaluation design method**

## **Focusing Sessions: A Stakeholder Engagement and Evaluation Design Method**

*A. Guidoccio<sup>1</sup>, V. Smith<sup>2</sup>*

<sup>1</sup> Ministry of Modernization- National Government Argentina, Director of Evaluation & Lead Policy Officer  
G20 DETF-Department of Innovation & Open Government, Buenos Aires, Argentina

<sup>2</sup> data2insight, Founder, Seattle, USA

When working with diverse communities, allowing time for stakeholder engagement is foundational to knowledge co-creation. A focusing session is a best practice that comes from theory-driven evaluation that contributes to better informed evaluation questions, which leads to more robust evaluation design, and findings that stakeholders value and use to inform decision making and action. It provides stakeholders with the opportunity to share with evaluators their goals, community and cultural context, understanding of their program, and desired impact. It also results in common understanding, and explicit, testable program theory, specific evaluation questions, and a framework for data collection and analysis. In this session participants will walk through the focusing session steps including program theory validation. They will leave with a basic understanding of how a focusing session works, as well as a list of tools and resources they can apply in their own evaluation practice to engage stakeholders and increase use of evaluation findings.

**RT 01 Focusing sessions: a stakeholder engagement and evaluation design method**

## **RC 01 - Round-table Contribution**

*A. Guidoccio<sup>1</sup>*

<sup>1</sup> Ministry of Modernization- National Government Argentina, Director of Evaluation & Lead Policy Officer  
G20 DETF-Department of Innovation & Open Government, Buenos Aires, Argentina

**RT 01 Focusing sessions: a stakeholder engagement and evaluation design method**

## **RC 02 - Round-table Contribution**

*V. Smith<sup>1</sup>*

<sup>1</sup> data2insight, Founder, Seattle, USA

Wednesday, 3 October 2018  
12:00–13:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**RT 02 When Focusing on Results Makes Impact Less Likely, and When it Doesn't**

## When Focusing on Results Makes Impact Less Likely, and When it Doesn't

*B. Douthwaite*<sup>1</sup>

<sup>1</sup> *Boru-Consult, Enabling and Evaluating Innovation in Agriculture, Westport, Ireland*

This round-table will examine when and how a results-focus makes impact more or less likely, and to identify and share principles, lessons, methods and strategies for monitoring and evaluation that supports the generation of beneficial outcomes. The first discussant (Boru Douthwaite) will argue that an increasing focus on achieving impact in research and development programmes makes impact less likely in complex and uncertain settings. In order to help secure funding, programme proponents – responding to the donor results agenda – make increasingly more heroic claims about expected impact, and, that this heroism is profoundly damaging. The second discussant (Marina Apgar) will argue that if programme proponents can keep the framing of results broad enough at the outset, and if funders are willing to take some risk, then using a results agenda could support increased impact. When the aim of a research for development programme is to address intractable challenges, such as is common in the large challenge-driven funding currently being promoted under the SDG agenda (such as the Global Challenges Research Fund in the UK), by design, programmes are aiming to build new and possibly unknown pathways to change. A moderator (Steff Deprez) will run subsequent discussion as a Fish Bowl (<http://www.kstoolkit.org/Fish+Bowl>) to encourage members of the audience to challenge or add to the propositions made by the two discussants, as well as share experiences of when M&E has enabled change. The Chair (Giel Ton) will introduce the topic and summarize the discussion at the end.

**RT 02 When Focusing on Results Makes Impact Less Likely, and When it Doesn't**

## RC 03 - When Focusing on Results Makes Impact Less Likely

*B. Douthwaite*<sup>1</sup>, *G. Ton*<sup>2</sup>, *M. Apgar*<sup>3</sup>, *S. Deprez*<sup>4</sup>

<sup>1</sup> *Boru-Consult, Principal, Westport, Ireland*

<sup>2</sup> *Institute of Development Studies, Centre for Impact Development, Brighton, United Kingdom*

<sup>3</sup> *Institute of Development Studies, Participation Research Cluster, Brighton, United Kingdom*

<sup>4</sup> *Independent Consultant, Principal, Antwerp, Belgium*

Increasing focus on achieving impact in research and development programmes makes impact less likely in complex and uncertain settings. In order to help secure funding, programme proponents – responding to the donor results agenda – make increasingly more heroic claims about expected impact, and, that this heroism is profoundly damaging.

It is damaging because it encourages the selection programmes with well-established innovation pathways because only well-known theories of change have any chance of plausibly linking research to impact at the outset of an intervention. Research that explores new or emerging innovation pathways is handicapped in the review process because it cannot provide a simple and known route to impact. Local-scale, bottom up development becomes devalued by impact inflation. The danger is, that, as the stream of outcomes from well-established trajectories dry up, the failure to encourage and support new, emerging and currently unknown trajectories will lead to a decrease in overall impact. Secondly, pathways that engage with the marginalized and excluded will not be explored and developed

**RT 02 When Focusing on Results Makes Impact Less Likely, and When it Doesn't****RC 04 - When Focusing on Results Could Make Impact More Likely**

*M. Apgar*<sup>1</sup>

<sup>1</sup> *Institute of Development Studies, Participation Research Cluster, Brighton, United Kingdom*

A results agenda can support increased impact if programme proponents can keep the framing of results broad enough at the outset, and if funders are willing to take some risk. When the aim of a research for development programme is to address intractable challenges, such as is common in the large challenge-driven funding currently being promoted under the SDG agenda (such as the Global Challenges Research Fund in the UK), by design, programmes are aiming to build new and possibly unknown pathways to change. For example, when programmes aim to embrace the 'leave no one behind' imperative of the SDGs, they have grapple with new and perhaps yet undiscovered pathways to overcome the poverty traps faced by the world's most excluded populations – slaves and bonded laborers, refugees and migrants, sexual minorities. In some cases, disruptive pathways may be more successful in creating real change and these are necessarily unknowable at the outset. In such cases, use of complexity-aware evaluation approaches that build on the experiences of marginalized people can enable programmes to learn their way to change – to see trajectories as they unfold, and thus catalyze real change when it starts to take shape and dampen undesired change. While aiming to reach the excluded is a normative starting point, pushing for a broader and more nuanced understanding of change can help to overcome the bias on linear and known pathways to impact in research for development programs.

**RT 02 When Focusing on Results Makes Impact Less Likely, and When it Doesn't****RC 05 - Moderator**

*B. Douthwaite*<sup>1</sup>, *G. Ton*<sup>2</sup>, *M. Apgar*<sup>3</sup>, *S. Deprez*<sup>4</sup>

<sup>1</sup> *Boru-Consult, Principal, Westport, Ireland*

<sup>2</sup> *Institute of Development Studies, Centre for Impact Development, Brighton, United Kingdom*

<sup>3</sup> *Institute of Development Studies, Participation Research Cluster, Brighton, United Kingdom*

<sup>4</sup> *Independent Consultant, Voices that Count, Antwerp, Belgium*

Steff will moderate the Fish Bowl (<http://www.kstoolkit.org/Fish+Bowl>) discussion to encourage the audience to take part in a round-table discussion. Anyone from the audience can join the discussants by sitting in one of three free chairs. When they have made their contribution they vacate the chair, making room for someone else to join. The discussion will be facilitated to encourage members of the audience to challenge or add to the propositions made by the two discussants. They will be particularly encouraged to share innovative methods in which they have used monitoring and evaluation in ways that embrace and work with complexity and uncertainty, and strategies they use to make and maintain space for them within mainstream evaluation practice.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

### RT 03 Evaluations in Fragile Contexts: How Do They Differ?

## Evaluations in Fragile Contexts: How Do They Differ?

*J. Pennarz<sup>1</sup>, R. Stryk<sup>2</sup>*

<sup>1</sup> IFAD, Independent Office of Evaluation, Rome, Italy

<sup>2</sup> UNRWA, Evaluation, Amman, Jordan

Working in fragile contexts is the current new reality for evaluators with fragility not confined to certain pockets or humanitarian contexts, but becoming increasingly widespread. Over 22% of the global population currently lives in fragile contexts, and this is anticipated to increase to 32% by 2050; the majority of the global poor – over 60% – will live in fragile contexts by 2030. The fluidity and diversity of fragile situations means that evaluators often cannot rely on tried-and-tested approaches. Evaluators have to adapt to the dynamics and complexity of the context; they have an even greater responsibility to contribute to a resilient society.

The aim of this Round Table is to raise and discuss some of the possible approaches, based on practical experiences from a range of different fragile contexts.

The speakers will address the following questions:

Fluidity. Is it appropriate to use standard evaluation criteria in situations that are as fluid as those “fragile situations”, and if so, how?

Fragility and related situations. Do evaluations in fragile, humanitarian, and post-conflict situations need to use a variety of lenses and how can they use conceptual frameworks to ensure an appropriate and relevant approach?

Fragility and gender. Do evaluators assess the effect of programs and policies on gender differently in such situations, and if so, how?

Link between self-evaluation and evaluation. Should evaluations rely on the internal perspectives gained through self-evaluation in order to understand the rapidly changing contexts characterising fragile situations, or are there alternatives?

Contributing to resilient societies. Can evaluation improve assessment of local risk and vulnerability for better and more targeted support, and thus contribute to the systematic understanding of how states and societies adapt and learn when faced with crises?

Responding to global commitments related to fragility. The international community has established guiding principles for country-led development and resilience-building in fragile situations, for example the Peace and State-building Goals (PSG) commitments under the New Deal. Therefore, should evaluations be more collaborative and country led? In fragile contexts, should evaluations assess the extent to which programs contribute to international commitments such as PSGs rather than just donor-specific measures?

Round Table process: The aim of the Round Table is to seek answers to the questions raised above. The composition of the round table has been carefully chosen. It will bring together evaluators from different contexts, representing professionals working within fragile contexts and those from international organisations conducting external evaluations. The expertise of the panelists covers a broad range of themes, including gender and rural poverty, humanitarian crisis, conflict and peace building, and development in fragile situations. The event will begin by a short presentation from IFAD narrating the experience from its own evaluations in different types of fragile or post-conflict situations, e.g. Palestine, Sri Lanka, Somalia and Georgia. This will be followed by a facilitated discussion, using the talk show format, where the discussants will be invited to respond to the questions above based on their own hands-on experiences.

**RT 03 Evaluations in Fragile Contexts: How Do They Differ?****RC 06 - Evaluations in Fragile Contexts: How Do They Differ?**

K. Abbot<sup>1</sup>

<sup>1</sup> USA

The international community has established guiding principles for country-led development and resilience-building in fragile situations, for example the Peace and State-building Goals (PSG) commitments under the New Deal. Therefore, should evaluations be more collaborative and country led? In fragile contexts, should evaluations assess the extent to which programs contribute to international commitments such as PSGs rather than just donor-specific measures?

**RT 03 Evaluations in Fragile Contexts: How Do They Differ?****RC 07 - Evaluations in Fragile Contexts: How Do They Differ?**

R. Stryk<sup>1</sup>

<sup>1</sup> UNRWA, Evaluation, Amman, Jordan

Fragility and related situations. Fragility and conflict share common root causes and reinforce each other; fragile countries tend to be prone to conflict while conflict can lead to a country being fragile or receiving such status. Ditto for humanitarian crisis. Do evaluations in such situations need to use a variety of lenses and how can they use conceptual frameworks to ensure an appropriate and relevant approach?

**RT 03 Evaluations in Fragile Contexts: How Do They Differ?****RC 08 - Evaluations in Fragile Contexts: How Do They Differ?**

K. Bitar<sup>1</sup>

<sup>1</sup> Palestinian Evaluation Association, Evaluation, Palestine, State of Palestine

Evaluations can miss out on some of the internal perspectives that lie with governments and NGOs, and that mark the continuously changing contexts characterising fragile situations. Initial design may not remain valid in rapidly shifting contexts, and evaluations need to be flexible accordingly. In such a scenario, should evaluations rely on the internal perspectives gained through self-evaluation in order to understand these dynamics?

**RT 03 Evaluations in Fragile Contexts: How Do They Differ?****RC 09 - Evaluations in Fragile Contexts: How Do They Differ?**

H. Khaira<sup>1</sup>

<sup>1</sup> IFAD, Independent Office of Evaluation, Rome, Italy

Delivery of the introductory presentation for the round table. The presentation will summarise the experience of the International Office of Evaluation of IFAD in conducting evaluations in fragile and conflict related situations. Three cases will be presented: Palestine, Sri Lanka and Georgia.

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

#### RT 04 Developing Synergies: the Evaluation Community and the Housing Sector

### Developing Synergies: the Evaluation Community and the Housing Sector

R. Meghir<sup>1</sup>, S. Edwards<sup>2</sup>

<sup>1</sup> Council of Europe Development Bank, Evaluation Department, Paris, France

<sup>2</sup> Housing Europe, the European Federation of Public, Cooperative and Social Housing, Brussels, Belgium

The right to adequate housing is recognized in many national constitutions as well as in the Universal Declaration of Human Rights. Housing is one of the central focal points of the sustainable development agenda that established, under SDG 11 (“Make cities and human settlements inclusive, safe, resilient and sustainable”), the target of “ensuring, by 2030, access, for all, to adequate, safe and affordable housing”. Inclusive and sustainable development of cities and urban settlements depends in fact on the capacity to provide a solution to the shortage of adequate and affordable housing. Access to housing may trigger processes of human empowerment, promote job opportunities and support economic development processes.

This roundtable focuses on two contemporary processes that involve the housing sector on the one hand, and the evaluation community on the other.

In recent years, the housing sector has increasingly gained ground in the international policy debate. Local, national and international actors have been called upon to help respond to the shortage of affordable housing through policies and programmes designed in a way that ensures social impact and sustainability. This has required the identification of suitable indicators for measuring housing deprivation and affordability; it has entailed investment in human and financial terms to sensitize and equip housing providers with tools to assess the social and environmental implications of their activities. At the same time, various initiatives are being tested at the local level to respond to housing needs of specific population groups (youth, refugees, low-income households.) in innovative ways. In this context, the evaluation community is called upon to generate, in a timely manner, knowledge on ‘what works’ by developing suitable and feasible methods for evaluation and impact assessment. Whereas a holistic approach is needed to correctly evaluate the performance of housing operations, there is often a tendency to focus on the timely realization and economic delivery of dwellings as proxies for determining the degree of success. The standard, internationally-recognized evaluation criteria (relevance, effectiveness, efficiency, impact and sustainability) need to be operationalized in light of the specific features of the housing sector. Additional methodological issues include the importance of correctly identifying objectives upfront and subsequently ensuring availability of the appropriate data to enable assessment of results and achievements. The right timing for evaluation of social results also needs to be determined.

**Format:** In the roundtable “Developing synergies: the evaluation community and the housing sector”, housing sector experts and evaluators will engage in a critical two-way dialogue on how evaluation knowledge could contribute to the work of housing stakeholders and, conversely, how to better capture housing issues in evaluation language and methodology. The duration of the roundtable is estimated at approximately 90 minutes and is structured around two blocks: (a) building blocks for rigorous, feasible and useful performance reporting by housing stakeholders; and (b) evaluating results of housing operations and their sustainability. Some of the issues expected to be tackled during the roundtable include: the role of key performance indicators; sensitization of housing stakeholders to evaluation principles within the context of their own corporate social responsibility; adequateness of the standard internationally-recognized evaluation criteria in providing a holistic assessment of performance of housing operations and how they can be more “customized” therefor.

**RT 04 Developing Synergies: the Evaluation Community and the Housing Sector**

**RC 10 - Round-table Contribution**

C. Heider<sup>1</sup>

<sup>1</sup> *The World Bank, Independent Evaluation Group, Washington DC, USA*

**RT 04 Developing Synergies: the Evaluation Community and the Housing Sector**

**RC 11 - Round-table Contribution**

R. Meghir<sup>1</sup>

<sup>1</sup> *Council of Europe Development Bank, Evaluation Department, Paris, France*

**RT 04 Developing Synergies: the Evaluation Community and the Housing Sector**

**RC 12 - Round-table Contribution**

L. Cuna<sup>1</sup>

<sup>1</sup> *Council of Europe Development Bank, Evaluation Department, Paris, France*

**RT 04 Developing Synergies: the Evaluation Community and the Housing Sector**

**RC 13 - Round-table Contribution**

S. Edwards<sup>1</sup>

<sup>1</sup> *Housing Europe, the European Federation of Public, Cooperative and Social Housing, Brussels, Belgium*

**RT 04 Developing Synergies: the Evaluation Community and the Housing Sector**

**RC 14 - Round-table Contribution**

G. Amitsis<sup>1</sup>

<sup>1</sup> *University of West Attica, Department of Business Administration, Athens, Greece*

**RT 04 Developing Synergies: the Evaluation Community and the Housing Sector**

**RC 15 - Round-table Contribution**

K. Scanlon<sup>1</sup>

<sup>1</sup> *London School of Economics, London, United Kingdom*

Wednesday, 3 October 2018  
14:30–16:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### RT 05 Digital – a Need to Have, not a Nice to Have in Evaluation

## Digital – a Need to Have, not a Nice to Have in Evaluation

*J. Robinson*<sup>1</sup>

<sup>1</sup> Kantar Public, International Development Practice, Washington, USA

Digital is no longer a nice-to-have but rather a need-to-have in the evaluator's toolbox. Digital practices connect people to each other as never before, and are transforming how policies and programmes are implemented. Industries of every kind have been impacted by this digital transformation, from communications channels to models of consumption and social interaction. Evaluation needs to reflect this newly dynamic and real-time environment by harnessing the opportunities that new, digital technology brings. We must keep pace with technological innovations by incorporating mobile methodologies into our practice, employing social media, and utilizing new observational and behavioral data sources to improve the depth and value of our research. This roundtable will kick-off discussion with a brief presentation on several ways we have incorporated digital technologies into evaluation work. These include more commonly used methods such as mobile data collection and real-time fieldwork dashboards as well as innovative methods such as eEthnography and focus groups using WhatsApp for richer qualitative data. For example, we are evaluating improvements to the Netherlands' transportation system through use of real-time mobile surveys with train riders, using sensors to measure hand washing habits in Ghana and road safety in the UK, and facilitating a WhatsApp group to track electoral opinions in South Africa. The ensuing discussion will solicit examples of how others are incorporating new technologies into evaluation work. Participants will also be invited to discuss challenges in accessing and using new technology, methodological drawbacks and tradeoffs, and integrating various digital data forms for meaningful analysis. We will also cover ethical concerns such as data security and responsible data practice in the increasingly digital age of evaluation. Good digital data practice guidelines are emerging from governments and agencies. These will be important for evaluators to watch as they design new digital approaches to evaluation. This roundtable will be accessible to novice and experienced evaluators alike, as well as evaluators with varying levels of comfort in accessing and using digital technology in their work.

### RT 05 Digital – a Need to Have, not a Nice to Have in Evaluation

## RC 16 - Round-table Contribution

*J. Robinson*<sup>1</sup>

<sup>1</sup> Kantar Public, International Development Practice, Washington, USA

Using sensors to measure hand washing habits in Ghana. Ms. Robinson will serve as chair for the roundtable. She will introduce the topic and the roundtable participants, present the example of sensor data collection in Ghana, and facilitate the ensuing roundtable discussion. She brings experience using digital data collection, SMS surveys, and fieldwork dashboards. (1/3 time).

**RT 05 Digital – a Need to Have, not a Nice to Have in Evaluation**

**RC 17 - Round-table Contribution**

G. Gault<sup>1</sup>

<sup>1</sup> Kantar Public, Digital Practice, Paris, France

Using WhatsApp to track electoral opinions. Ms. Gault will serve as our lead expert in digital. She will present the example of using WhatsApp to track electoral opinions and lead the response to questions from roundtable participants who are using or considering digital approaches. (1/3 time).

**RT 05 Digital – a Need to Have, not a Nice to Have in Evaluation**

**RC 18 - Round-table Contribution**

D. Ulicna<sup>1</sup>

<sup>1</sup> Kantar Public, Brussels, Belgium

Real-time mobile surveys in the transport sector in the Netherlands. Ms. Ulicna will present work using in-the-moment research. She brings experience using digital data collection, SMS surveys, and fieldwork dashboards, as well as text mining software. (1/3 time).

Wednesday, 3 October 2018  
16:30–18:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**RT 06 Institutionalization of Evaluation in Europe: From Boom Times to New Turbulences?**

## **Institutionalization of Evaluation in Europe: From Boom Times to New Turbulences?**

*W. Meyer<sup>1</sup>, P.D.R. Stockmann<sup>1</sup>, T. Widmer<sup>2</sup>, P. Dahler-Larsen<sup>3</sup>, M. Bustelo Ruesta<sup>4</sup>, T.D. Iacob<sup>5</sup>, J. Remr<sup>6</sup>, L. Taube<sup>1</sup>*

<sup>1</sup> Saarland University, CEval, Saarbrücken, Germany

<sup>2</sup> University of Zurich, Department of Political Science, Zurich, Switzerland

<sup>3</sup> University of Copenhagen, Department of Political Science- Institut for Statskundskab, København K, Denmark

<sup>4</sup> Universidad Complutense Madrid, Departamento de Ciencia Política y de la Administración, Pozuelo de Alarcón Madrid, Spain

<sup>5</sup> National School of Political and Administrative Studies Bucharest, Departamentul de Relații Internaționale și Integrare Europeană, Bucharest, Romania

<sup>6</sup> Univerzity Karlovy/Institute for Evaluations and Social Analyses - INESAN, Fakulta sociálních, Prague 7, Czechia

This round-table discussion is about the results of the GLOBE project on the Institutionalization of Evaluation in Europe. Started at the EES-Conference in Maastricht two years ago, the European part of this project is about to be finalized in a book supposed to be published in 2019. Authors from 16 European countries investigated the state of institutionalization of evaluation by following a shared analysis guideline. The editors Reinhard Stockmann and Wolfgang Meyer from the Centre for Evaluation CEval at Saarland University will present the results of their comparative analysis for the first time. Some of the authors will comment 'critically from a national perspective and challenge these results. The main task of this session is to adjust and verify the comparative analysis, giving the authors and the participants an opportunity to contribute to the analytical process. During the session, the round-table tries to explore the similarities and differences of the institutionalization processes in Europe. Is there a commonly shared development with the same, comparable drivers of the process in Europe? Which role do national traditions, differences in the political and social systems, the scientific culture, the focal points of evaluation and other drivers play? Is there a pressure for coherence from international organisations, especially the EU? Or is evaluation culture decomposed by divergent developments in various policy departments with their own rules, institutions and traditions? Which challenges does evaluation face in Europe – both from the perspective of a joint European culture and the diversity of national political demands? And finally: are there any recommendations how the Evaluation community in Europe in general and the EES in particular should act to improve the institutionalization of evaluation in Europe? The session consists of four parts. First, Reinhard Stockmann will give a short introduction on the comparative findings and offer some thesis on the questions above. Second, some of the authors (appr. 4–5) will reply to these thesis from a national perspective, highlighting the national specifics and the main results of their own analysis in contrast to the shared European perspective. Third, the other participants of the round table will get the opportunity to discuss these contributions. Finally, the floor will be opened for the participants to share their thoughts on the discussed issues. The results of all these discussions are noted and summarized in the end, being used for revising the comparative chapter and its final conclusions.

**RT 06 Institutionalization of Evaluation in Europe: From Boom Times to New Turbulences?**

**RC 19 - Round-table Contribution**

*W. Meyer<sup>1</sup>, P.D.R. Stockmann<sup>1</sup>, T. Widmer<sup>2</sup>, P. Dahler-Larsen<sup>3</sup>, M. Bustelo Ruesta<sup>4</sup>, T.D. Iacob<sup>5</sup>, J. Remr<sup>6</sup>, L. Taube<sup>1</sup>*

<sup>1</sup> Saarland University, CEval, Saarbrücken, Germany

<sup>2</sup> University of Zurich, Department of Political Science, Zurich, Switzerland

<sup>3</sup> University of Copenhagen, Department of Political Science- Institut for Statskundskab, København K, Denmark

<sup>4</sup> Universidad Complutense Madrid, Departamento de Ciencia Política y de la Administración, Pozuelo de Alarcón Madrid, Spain

<sup>5</sup> National School of Political and Administrative Studies Bucharest, Departamentul de Relații Internaționale și Integrare Europeană, Bucharest, Romania

<sup>6</sup> Univerzity Karlovy/Institute for Evaluations and Social Analyses - INESAN, Fakultu sociálních, Prague 7, Czechia

**RT 06 Institutionalization of Evaluation in Europe: From Boom Times to New Turbulences?**

**RC 20 - Round-table Contribution**

*W. Meyer<sup>1</sup>, P.D.R. Stockmann<sup>1</sup>, T. Widmer<sup>2</sup>, P. Dahler-Larsen<sup>3</sup>, M. Bustelo Ruesta<sup>4</sup>, T.D. Iacob<sup>5</sup>, J. Remr<sup>6</sup>, L. Taube<sup>1</sup>*

<sup>1</sup> Saarland University, CEval, Saarbrücken, Germany

<sup>2</sup> University of Zurich, Department of Political Science, Zurich, Switzerland

<sup>3</sup> University of Copenhagen, Department of Political Science- Institut for Statskundskab, København K, Denmark

<sup>4</sup> Universidad Complutense Madrid, Departamento de Ciencia Política y de la Administración, Pozuelo de Alarcón Madrid, Spain

<sup>5</sup> National School of Political and Administrative Studies Bucharest, Departamentul de Relații Internaționale și Integrare Europeană, Bucharest, Romania

<sup>6</sup> Univerzity Karlovy/Institute for Evaluations and Social Analyses - INESAN, Fakultu sociálních, Prague 7, Czechia

This round-table discussion is organized by the Centre for Evaluation CEval at Saarland University. The session is chaired by Adj. Prof. Dr. Wolfgang Meyer, Vice-director of CEval and senior researcher at Saarland University. He is working for more than 20 years on Evaluations, merely on development cooperation, environmental and labour market issues. The main contributions are made by authors of the GLOBE-project on the Institutionalization of Evaluation in Europe. They are all highly respected and experience evaluators from different European countries, some of them are engaged in the European Evaluation Society for years.

**RT 06 Institutionalization of Evaluation in Europe: From Boom Times to New Turbulences?**

**RC 21 - Round-table Contribution**

*W. Meyer<sup>1</sup>, P.D.R. Stockmann<sup>1</sup>, T. Widmer<sup>2</sup>, P. Dahler-Larsen<sup>3</sup>, M. Bustelo Ruesta<sup>4</sup>, T.D. Iacob<sup>5</sup>, J. Remr<sup>6</sup>, L. Taube<sup>1</sup>*

<sup>1</sup> Saarland University, CEval, Saarbrücken, Germany

<sup>2</sup> University of Zurich, Department of Political Science, Zurich, Switzerland

<sup>3</sup> University of Copenhagen, Department of Political Science- Institut for Statskundskab, København K, Denmark

<sup>4</sup> Universidad Complutense Madrid, Departamento de Ciencia Política y de la Administración, Pozuelo de Alarcón Madrid, Spain

<sup>5</sup> National School of Political and Administrative Studies Bucharest, Departamentul de Relații Internaționale și Integrare Europeană, Bucharest, Romania

<sup>6</sup> Univerzity Karlovy/Institute for Evaluations and Social Analyses - INESAN, Fakultu sociálních, Prague 7, Czechia

**RT 06 Institutionalization of Evaluation in Europe: From Boom Times to New Turbulences?**

**RC 22 - Round-table Contribution**

*W. Meyer<sup>1</sup>, P.D.R. Stockmann<sup>1</sup>, T. Widmer<sup>2</sup>, P. Dahler-Larsen<sup>3</sup>, M. Bustelo Ruesta<sup>4</sup>, T.D. Iacob<sup>5</sup>, J. Remr<sup>6</sup>, L. Taube<sup>1</sup>*

<sup>1</sup> Saarland University, CEval, Saarbrücken, Germany

<sup>2</sup> University of Zurich, Department of Political Science, Zurich, Switzerland

<sup>3</sup> University of Copenhagen, Department of Political Science- Institut for Statskundskab, København K, Denmark

<sup>4</sup> Universidad Complutense Madrid, Departamento de Ciencia Política y de la Administración, Pozuelo de Alarcón Madrid, Spain

<sup>5</sup> National School of Political and Administrative Studies Bucharest, Departamentul de Relații Internaționale și Integrare Europeană, Bucharest, Romania

<sup>6</sup> Univerzity Karlovy/Institute for Evaluations and Social Analyses - INESAN, Fakultu sociálních, Prague 7, Czechia

Wednesday, 3 October 2018  
16:30–18:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**RT 07 How Is Resilience Measurement Moving us Beyond Just Good M&E?: The State of Resilience Evaluation Practice and the Paths Ahead**

## How Is Resilience Measurement Moving us Beyond Just Good M&E?: The State of Resilience Evaluation Practice and the Paths Ahead

*B. Sagara*<sup>1</sup>

<sup>1</sup> Mercy Corps, Deputy Director for Research and Learning, Washington- DC, USA

The concept of resilience is on a journey from theoretical niche towards widespread application. It has challenged measurement and evaluation specialists across sectors, disciplines and geographies. Evaluation of resilience demands many of the innovations needed in the evaluation field more broadly, including evidence to inform adaptive management, practical applications of complexity theory and systems thinking, real-time evaluation to inform high frequency decision making, analysis of the multiple trajectories of development, and understanding the impacts of integrated, multi-stakeholder interventions. A wide range of frameworks and approaches are emerging from diverse set of actors and sectors to diagnose, measure and evaluate resilience. These include evaluation methods of resilience to food security shocks, urban resilience, resilience of socio-ecological systems, and individual psychosocial resilience, to name a few. The approaches applied across these populations and contexts share some common principles and features, such as an emphasis on analyzing exposure to or recovery from shocks and stressors, and attention to non-linear relationships. However, there also remain major gaps between the methodological approaches to resilience measurement and evaluation that have emerged out of different disciplines and sectors. Efforts to understand and bridge these differences are critical to advancing the field of evaluation of resilience, and to generating robust evidence and learning on the efficacy of various investments to strengthen the resilience of societies and systems.

This roundtable will take stock of the field of resilience measurement, evaluation and learning, and solicit examples and insights from experts to identify promising practices and frontier challenges and opportunities. The objectives of the round table are:

To present the current state of resilience measurement and discuss how the evaluation field has already risen to the unique challenges posed by the concept of resilience, including highlighting examples of innovative evaluation practice; To discuss the state of resilience measurement across and between sectors, and draw out common principles and practices, as well as identify where we need to forge stronger connections across disciplines to maximize learning; and To propose a set of critical challenges and opportunities for the evaluation field in regard to resilience in the coming 5 years, which will help define a shared learning agenda. This session will be grounded in practical examples. Experts involved in the Resilience Measurement, Evaluation, and Learning Community of Practice will draw on their experiences in developing and applying innovative approaches to evaluating programs and policies aimed at strengthening resilience of at-risk and crisis-affected populations. The session will be relevant to donor representatives and practitioners seeking to better understand the state of the art and frontier questions regarding approaches monitoring and evaluating investments in resilience.

The format for the roundtable will be a conversational debate style, with a dynamically moderated group of discussants. Following opening remarks by the chair and each speaker, participants will have an opportunity to pose questions with the aim of clarifying, challenging or affirming the key points made. This interactive format will build engagement and understanding of the primary issues and question to be address in the session.

**RT 07 How Is Resilience Measurement Moving us Beyond Just Good M&E?: The State of Resilience Evaluation Practice and the Paths Ahead****RC 24 - Evaluating Large-Scale Climate Resilience Investments: What Does Complexity Theory Bring?**

J. Puri<sup>1</sup>

<sup>1</sup> Green Climate Fund, Independent Evaluation Unit, Yeonsu-gu- Incheon 406-840, Republic of Korea

Dr Puri's remarks will contribute to the overall objectives described in the overview of this round table session (see main abstract for details). In particular, she will lead a discussion of the state of resilience measurement across and between sectors, and draw out common principles and practices, as well as identify where we need to forge stronger connections across disciplines to maximize learning.

**RT 07 How Is Resilience Measurement Moving us Beyond Just Good M&E?: The State of Resilience Evaluation Practice and the Paths Ahead****RC 25 - Resilience Measurement: Major Achievements and Outstanding Challenges**

T. Frankenberger<sup>1</sup>

<sup>1</sup> TANGO International, President, Tuscon- AZ, USA

Dr Frankenberger's remarks will contribute to the overall objectives of the session as outlined in the abstract overview. In particular, he will present the current state of resilience measurement and discuss how the evaluation field has already risen to the unique challenges posed by the concept of resilience, including highlighting examples of innovative evaluation practice.

**RT 07 How Is Resilience Measurement Moving us Beyond Just Good M&E?: The State of Resilience Evaluation Practice and the Paths Ahead****RC 26 - Evaluating Resilience in Urban Areas and Populations**

M. Khan<sup>1</sup>

<sup>1</sup> Daira, Independent Consultant, London, United Kingdom

Dr Khan will contribute to the overall objectives of the session as outlined in the abstract overview. In particular, she will lead a discussion on the critical challenges and opportunities for the evaluation field in regard to resilience in the coming 5 years, which will help define a shared learning agenda.

**RT 07 How Is Resilience Measurement Moving us Beyond Just Good M&E?: The State of Resilience Evaluation Practice and the Paths Ahead****RC 27 - Round-table Contribution**

C. Hutchings<sup>1</sup>

<sup>1</sup> Oxfam Great Britain, Program Quality, Oxford, United Kingdom

Wednesday, 3 October 2018  
16:30–18:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**RT 08 Joint Session of the TWG on Private Sector Evaluation (PSE) and the AEA TIG on Social Impact Measurement (SIM)**

## Joint Session of the TWG on Private Sector Evaluation (PSE) and the AEA TIG on Social Impact Measurement (SIM)

*F. Korfker*<sup>1</sup>

<sup>1</sup> Independent consultant former EBRD, n.a., Amsterdam, Netherlands

**Rationale International Cooperation:** To explain the content of the TWG on Private Sector Evaluation and to introduce the TIG SIM to the EES members and learn about their way of working. Learning from each other and exploring opportunities to collaborate on topics of joint interest will be highlighted.

### Objectives sought:

**Objective A:** To highlight details of the working of the four Sub-Groups of the TWG on PSE to interest EES members to join the Group: 1. Sub-group on the Evaluation of Financial Intermediaries (including SMEs) 2. Sub-group on the Evaluation of Public-Private Partnerships (PPPs) 3. Sub-group on Corporate Self-Evaluation and Social Impact investing 4. Sub-group on Specificity and Dynamics of Private Sector Evaluation Presentation 1: Fredrik Korfker, leader of the TWG on PSE: activities of the TWG over the past two years and what the Group intends to accomplish in the future; Presentation 2: Mehmet Uzunkaya presents PPP Evaluation research that his Sub-group 2 conducted;

**Objective B:** Learn about the working of AEA's TIG on social impact measurement, its membership and key activities and identify areas where the AEA and EES can collaborate through interaction of both Groups. Presentation 3: Veronica Olazabal will highlight how the TIG SIM is working, what the main objectives are and what kind of activities have been carried out over the past year including in partnership with the Regional Evaluation Associations (e.g., AfrEA and the South East Asian Conclave). Presentation 4: Kruno Karlovcec wants to elaborate on the social impact in a bottom-up sense, f.e. "from local to global" or reframed "from community entrepreneur to impact investor". Data would be derived from so far supply and demand of services and assistance to (social) entrepreneurs at the one-stop-shops. Presentation 5: Sarah Olson (to be confirmed) on social impact measurement: state of the art, challenges and opportunities in respect of the Sustainable Development Goals. Justification: As TWGs are stimulated by the EES Board, a promotion session of the TWG on Private Sector Evaluation is important. It is also important to introduce the TIG on social impact measurement as, so far, the level of collaboration between the TWGs and TIGs has been limited. Both, the TIG SIM as well as the TWG PSE deal with evaluation of private sector development interventions which is crucial for holding the stakeholders to account. It is essential to explore the critical role of evaluation in private sector economic activity. Private sector economic activities have tremendous impacts on the planet's ability to sustain human life, human rights, and also represents a capacity to generate helpful innovation and deliver it at massive scale, there is an urgent need for private sector actors to understand the potential and actual impact of their economic activities. Today this need is recognized more clearly than ever before. As such this is the time for global coordination to accelerate advances in and adoption of PSE practice and the interaction between the EES and the AEA at the level of the TWG and TIG can be important in this respect.

**RT 08 Joint Session of the TWG on Private Sector Evaluation (PSE) and the AEA TIG on Social Impact Measurement (SIM)****RC 30 - Thematic Working Group on Private Sector Evaluation**

*F. Korfker*<sup>1</sup>

<sup>1</sup> *Independent consultant former EBRD, n.a., Amsterdam, Netherlands*

Presenting the activities of the TWG over the past two years and what the Group intends to accomplish in the future (10 minutes).

**RT 08 Joint Session of the TWG on Private Sector Evaluation (PSE) and the AEA TIG on Social Impact Measurement (SIM)****RC 31 - Evaluation of PPPs**

*M. Uzunkaya*<sup>1</sup>

<sup>1</sup> *Middle East Technical University, Ankara, Turkey*

Review of PPP Evaluation research that the PPP TWG Sub-Group has carried out over the past two years (10 minutes).

**RT 08 Joint Session of the TWG on Private Sector Evaluation (PSE) and the AEA TIG on Social Impact Measurement (SIM)****RC 32 - Presentation of the Topical Interest Group(TIG) on Social Impact Measurement of the American Evaluation Association (AEA)**

*V. Olazabal*<sup>1</sup>

<sup>1</sup> *The Rockefeller Foundation, Director of Measurement- Evaluation and Organizational Performance, City, Netherlands*

Highlight how the TIG SIM is working, what the main objectives are and what kind of activities have been carried out over the past year including in partnership with the Regional Evaluation Associations (e.g., AfrEA and the South East Asian Conclave). (10 minutes).

**RT 08 Joint Session of the TWG on Private Sector Evaluation (PSE) and the AEA TIG on Social Impact Measurement (SIM)****RC 33 - Utilizing Bottom-Up Approach in Social Impact: Going from Community Entrepreneur to Impact Investor**

*K. Karlovcec*<sup>1</sup>

<sup>1</sup> *Ministry of Economic Development and Technology, Entrepreneurship Unit, Ljubljana, Slovenija*

He will elaborate on the social impact by utilizing a bottom-up approach, that is, going "from local to global" or "from community entrepreneur to impact investor". Data will be derived from the supply and demand of services and assistance to (social) entrepreneurs at the Slovenian one-stop-shops.

**RT 08 Joint Session of the TWG on Private Sector Evaluation (PSE) and the AEA TIG on Social Impact Measurement (SIM)**

**RC 34 - Messages from the Social Impact Measurement Arena**

*M. Harnar<sup>1</sup>*

<sup>1</sup> *Western Michigan University, Assistant Professor of Interdisciplinary Evaluation, USA*

<sup>2</sup> *Social Value US, Board member, USA*

Provide insight into the current Social Value US strategic direction, highlighting trainings and coalitions focused on social impact management and analysis.

Wednesday, 3 October 2018  
16:30–18:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**RT 09 The Road Ahead for Evaluation in the Lusophone World: Resilience and Action in Critical Times**

## The Road Ahead for Evaluation in the Lusophone World: Resilience and Action in Critical Times

*E. Sarmento<sup>1</sup>, M. Branco<sup>2</sup>*

<sup>1</sup> Nova School of Business and Economics, Novafrica, Lisbon, Portugal

<sup>2</sup> Independent Consultant, EES Board- Portugal Impact Group, Porto, Portugal

**Rationale:** When it comes to the area of monitoring and evaluation, Lusophone countries are lagging behind, as the initiatives to establish a common culture of evaluation practice among Portuguese-speaking countries and communities are still incipient.

This panel seeks to contribute to a better understanding and increased awareness of the need to foster evaluation practices within the unique political and cultural setting of the Portuguese-speaking world. Each individual presentation will have a 10 minutes duration.

**Justification and brief narrative:** The Community of Portuguese Language Countries (Comunidade dos Países de Língua Portuguesa or CPLP) grew from 7 countries, Angola, Brazil, Cabo Verde, Guinea-Bissau, Mozambique, Portugal, and São Tomé and Príncipe to the current 9, after the self-determination of Timor-Leste in 2002 and the accession of Equatorial Guinea in 2014. Portuguese is the sixth most natively spoken language in the world and Portuguese-speaking countries are home to 267 million people located in four continents who have a common language, a shared history, and cultural similarities.

Evaluation is of critical importance for policy makers' decision processes, underlying several countries' strategies for accelerated growth and sustainable development worldwide. However, in Lusophone-speaking countries the field of evaluation is still fragmented, with a lack of strong professional organizations dedicated to evaluation, specialized evaluation journals and overall, a lack of capacity to conduct advocacy and influence policy-making. Many political decisions are still driven by values rather than outcomes, compounded by the lack of an evaluation culture and the set of skills for using rigorous evidence, and the mismatch between political timetables and the timelines of the evidence producers.

The scarce cases of dedicated academic training (e.g. Masters in evaluation in specialized fields in Brazil) and well established and proactive communities of practice has not done enough towards the diffusion of an evaluation culture and the development of a professional class able to promote and disseminate learning and evaluation research.

Despite some initiatives in Portugal and Brazil, where sectorial and national evaluation societies, along with NGOs have emerged, other Lusophone-speaking countries are lagging even more behind, seeing mostly small and informal networks fostering short-term training activities without a long-term view.

**Objectives:** In close collaboration with representatives of several Lusophone countries, this panel intends to:

Map the background, governance structure of evaluation communities and activities in the different Lusophone countries;

Discuss the integration of evaluation as an essential component of public policy, programs and projects in Portuguese speaking countries, unbundling the concept of "Lusophone world" from an evaluation perspective;

Disseminate new initiatives and evaluation activities involving Lusophone countries;

Discuss the kind of support and practical arrangements needed by Lusophone countries to engage further with evaluation, with a focus on supply and demand triggers and barriers standing in the way of more systematic use of evidence and evaluation;

Discuss future actions and partnerships for the development of evaluation capabilities, to strengthen impact evaluation capacities of national actors and networks and create synergies with researchers and implementing agents to produce more credible evaluation results.

**RT 09 The Road Ahead for Evaluation in the Lusophone World: Resilience and Action in Critical Times**

**RC 35 - AfDB's Evaluations in Lusophone countries in Africa: Experience, Challenges and Opportunities on the Way Forward**

C. Félix<sup>1</sup>

<sup>1</sup> African Development Bank, Idev, Lisbon, Portugal

Discuss the kind of support and practical arrangements needed by Lusophone countries to engage further with evaluation, with a focus on supply and demand triggers and barriers standing in the way of more systematic use of evidence and evaluation – the experience of an evaluator working with IDEV/ AfDB; Reflect on the opportunities, new initiatives and evaluation activities involving Lusophone countries that can be taken ahead by multilateral development Banks.

**RT 09 The Road Ahead for Evaluation in the Lusophone World: Resilience and Action in Critical Times**

**RC 36 - What Can We Learn from the Institutionalization of M&E in Brazil?**

GL. Weber Costa<sup>1</sup>

<sup>1</sup> CLEAR, CLEAR Brazil, São Paulo, Brazil

- Map the background, governance structure: Brazil
- Discuss the integration of evaluation as an essential component of public policy, programs and projects in Portuguese speaking countries, unbundling the concept of “Lusophone world” from an evaluation perspective;
- Discuss the kind of support and practical arrangements needed by Lusophone countries to engage further with evaluation, with a focus on supply and demand triggers and barriers standing in the way of more systematic use of evidence and evaluation;
- Disseminate new initiatives and evaluation activities involving Lusophone countries.

**RT 09 The Road Ahead for Evaluation in the Lusophone World: Resilience and Action in Critical Times**

**RC 37 - Building Evaluation Capacity: Early Lessons from the WACIE Initiative**

M. Gaarder<sup>1</sup>

<sup>1</sup> 3ie, Director, Oslo, Norway

- 3ie, with resources from Hewlett, is carrying out an assessment of the state of institutionalized evaluation capacity in 8 (mostly) francophone West African countries. The study will identify existing resources as well as barriers and needs. The talk will focus on early lessons relevant to the lusophone context;
- Discuss future actions and partnerships for the development of evaluation capabilities, to strengthen impact evaluation capacities of national actors and networks and create synergies with researchers and implementing agents to produce more credible evaluation results;
- Disseminate new initiatives and evaluation activities involving Lusophone countries.

**RT 09 The Road Ahead for Evaluation in the Lusophone World: Resilience and Action in Critical Times**

**RC 38 - Present and Future Challenges of the Evaluation in Portugal**

R. Godinho M. C.<sup>1</sup>

<sup>1</sup> IESE, Professor, Lisbon, Portugal

- Map the background, governance structure: Portugal.
- Reflect on the present and future challenges of the evaluation in Portugal, namely the sustainability of external evaluators, uses of evaluations, areas of evaluation and the implementation of an evaluation network in Portuguese-speaking countries;
- Discuss the profile, role and skills of the evaluator in the future.

**RT 09 The Road Ahead for Evaluation in the Lusophone World: Resilience and Action in Critical Times**

**RC 39 - The Policy-Maker Perspective: What Are the Bottlenecks to Evaluation in Cabo Verde and Africa**

V. Borges<sup>1</sup>

<sup>1</sup> Cabo Verde Foundation for Development and International Engagement, Fdi-Cv, Praia, Cabo Verde

- Map the background, governance structure: Cabo Verde, recalling the diversity of lusophone countries in terms of size, population, natural resources, level of economic, social and human development, democracy and governance;
- Discuss the integration of evaluation as an essential component of public policy, programs and projects in Portuguese speaking countries, unbundling the concept of "Lusophone world" from an evaluation perspective;
- Discuss the role and responsibility of national governments, civil society organizations, donors and development partners in developing evaluation capacity, under the topic of "endogenous evaluation";
- Disseminate new initiatives and evaluation activities involving Lusophone countries.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**RT 10 Scrambling after Moving Targets: M&E Applied to Adaptive Management in Peacebuilding**

## Scrambling after Moving Targets: M&E Applied to Adaptive Management in Peacebuilding

*L. Jean<sup>1</sup>, E. Forsyth Queen<sup>2</sup>*

<sup>1</sup> CDA Collaborative Learning, Collaborative Learning and Evaluation, Cambridge, USA

<sup>2</sup> Brandeis University, Heller School for Social Policy and Management- Program on Conflict and Coexistence, Waltham- MA, USA

This roundtable will focus on practical aspects of using evaluation to inform adaptive management in FCAS. We will share practical lessons and tools tested in recent evaluative engagements in FCAS and our research with practitioners. The session will examine the nexus of adaptive management and evaluation by drawing on participant experiences and reflections.

We see adaptive peacebuilding programming as analogous to sailing a boat – in contrast to laying railroad tracks and driving a train. Trains proceed along a linear pathway, arriving and departing from fixed stations on a predictable schedule (although delays and accidents do occur!). Classic development programming appears much like the railroad system, requiring identification of clear mileposts, a timeline and an ultimate attainable objective, as laid out in the logframe, Gantt chart, and M&E plan. Evaluation, in most cases, is relatively straightforward: has the program achieved its intermediate objectives and contributed to a longer-term goal on schedule and within budget parameters?

In contrast, adaptive management – in development and peacebuilding – is analogous to sailing a boat, which requires constant response to changing conditions of wind, current, unexpected obstacles in the waterway, the effects of fog and storms, the condition of the boat, and the variable performance of the crew. In most circumstances, particularly when sailing upwind, making progress also involves frequent changes in direction through “tacking,” although the ultimate objective remains fixed. Under adaptive management, a peacebuilding program initiative or intervention could consider a wide range of changes: activities, participants, timing, resources, scale, geographic focus, linkages with other programming, and staffing. Even intermediate objectives and longer-term goals can be reconsidered, if conditions warrant. In addition, peacebuilding programs operating under adaptive management need to remain open to evidence that the preliminary Theory of Change is wrong in the whole or in part. With so many variables in play and subject to change, evaluation becomes more difficult!

Our research shows that where adaptive management has been applied in peacebuilding, it is poorly documented and poorly reported as part of standard reporting templates. Much of the adaptive nature of operational decisions at the field level is lost in the proverbial “black box of decision-making,” and program adjustments are difficult to reconstruct retrospectively. Also, in big donor bureaucracies, there may be pockets of enthusiastic champions of adaptive approaches, but there are still many staff wedded to the more traditional – and familiar – approaches embedded in results-based management and rigid M&E frameworks. However, in complex and turbulent environments, we cannot afford to overlook emerging results and critical decisions taken to course correct. Effective use of M&E to inform adaptive management is crucial if we aim to support more resilient societies. How can adaptive management bridge gaps between evaluation systems and changes in programming and strategies? How can evaluators design and manage evaluations to pave the way for adaptive management in unpredictable contexts? How can evaluators influence decision-makers about required course corrections?

This session invites a reality check and an honest reflection on the value and opportunity that adaptive management offers to support effective and strategic programming in FCAS.

**RT 10 Scrambling after Moving Targets: M&E Applied to Adaptive Management in Peacebuilding****RC 40 - Connect. Learn. Adapt. Repeat. Using Adaptive Management to Bolster Evaluation and Resiliency in Conflict-Affected Areas**

*E. Forsyth Queen*<sup>1</sup>

<sup>1</sup> *Independent Consultant, Boston, USA*

Emily Forsyth Queen, Heller School for Social Policy and Management, Brandeis University

In complex and turbulent environments, the effective use of monitoring and evaluation to inform adaptive management is crucial if we aim to support more resilient societies. This session will explore challenges and opportunities for improving adaptive management practice using findings from semi-structured interviews with evaluators and program staff working on peacebuilding and development in conflict-affected areas. Roundtable discussion will focus on: How can adaptive management bridge gaps between evaluation systems and changes in programming and strategies? How can evaluators design and manage evaluations to pave the way for adaptive management in unpredictable contexts? Who should be considered an evaluator when centering voices of those impacted by programs and their ideas on how course corrections can lead to more resiliency? How can evaluators influence those with more power – especially donors – about changing programs, processes, and policies to better use data and support adaptive management?

**RT 10 Scrambling after Moving Targets: M&E Applied to Adaptive Management in Peacebuilding****RC 41 - Existing Evaluation Approaches and How They Fit with Adaptive Management**

*I. Jean*<sup>1</sup>

<sup>1</sup> *CDA Collaborative Learning, Collaborative Learning and Evaluation, Cambridge, USA*

Existing Evaluation Approaches and How They Fit with Adaptive Management Isabella Jean, Director, CDA Collaborative Learning Projects Peacebuilding practitioners often repeat the observation that conflict contexts are fluid, rapidly changing and uncertain. If we accept the reality of constant change as true, then peace initiatives of all kinds are best designed and implemented in accordance with adaptive management principles. In the peacebuilding context, adaptive management, done well, leads to deeper understanding of how a conflict system functions and constant assessment of what kinds of initiatives are working – which enables more effective generation of desired changes. Our experience shows that systems thinking tools are also well suited to developing a more nuanced understanding of complex interactions and for obtaining ongoing feedback. At the same time, these approaches raise challenges for monitoring and evaluation processes in common use. Evaluations of adaptive management programs must consider the decision nodes – the determination of when to “tack” in our sailing analogy [see roundtable abstract]. Why did the program decide to change direction, or reconsider whether it is still reasonable to reach a certain destination? What was the information available? What were the observed short-term effects of the programming? What are alternative explanations for those effects? Were there unexpected results? What factors in the environment (cataclysmic political or security events, natural disasters...) necessitated a change of plans? While some of those questions are also applicable to classic approaches, they become more important when evaluating a program operating under adaptive management. M&E systems designed to support adaptive management should aid in documentation of decision making and the rationale for adjustments, as well as actual observable changes in the system, both expected and unexpected. What are existing evaluation methods that are appropriate for this kind of inquiry? We'd propose that Most Significant Change, Outcome Harvesting, and Developmental Evaluation can be considered “utilization-focused” (that is, striving to meet the needs of the implementing team/organization and aiming to generate applicable lessons) and relatively “goal-free” (that is, casting a wide net to capture changes, not focusing solely on the pre-determined goals and objectives in a program proposal). These methods can also be used

in combination. This roundtable session seeks to foster discussion and input from experienced EES participants on their use of these methods and observations on how these support (or not) adaptation and improvements. CDA's forthcoming article argues that adaptive management approaches – when combined with systems thinking tools for deeper understanding – show promise for provoking more sophisticated and effective peacebuilding initiatives. Before taking full advantage of these emerging methods, peacebuilding practitioners have some work to do – and the results will be well worth the time and resources devoted. As more initiatives undertake an adaptive approach, we must, as a field, pay careful attention to the successes, failures and lessons from the monitoring and evaluation methods applied.

Wednesday, 3 October 2018  
18:15–19:45

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**RT 11 At the Axiological Intersection of Epistemology and Ontology: Or, Where the Rubber Hits the Road**

## **At the Axiological Intersection of Epistemology and Ontology: Or, Where the Rubber Hits the Road**

*J. Whynot<sup>1</sup>, S. Montague<sup>2</sup>, J. Sylvestre<sup>3</sup>, R. Shepherd<sup>2</sup>*

<sup>1</sup> University of Ottawa, Student, Ottawa, Canada

<sup>2</sup> Carleton University, School of Policy and Public Administration, Ottawa, Canada

<sup>3</sup> University of Ottawa, Faculty of Social Sciences, Ottawa, Canada

This panel addresses Stream 2 of the conference theme “Rethinking evaluation methods and methodology”. The panel is comprised of a collective of evaluation academics including Deans, professors, and instructors who are consciously addressing the learning needs of the next generation of emerging evaluators. In North America, more than 48 university-based evaluation training programs are offered (LaVelle & Donaldson, 2010). A significant number of evaluators sidestep into the profession through sectors such as health, sociology and education (Stevahn, King, Ghere, & Minnema, 2005) and consequently do not hold advanced evaluation degrees and are motivated to seek advanced formal training (Christie, Quiñones, & Fierro, 2014). While these university programs are grounded in evaluation theory, a gap exists in bridging theoretical concepts to practical applications in evaluation work. Several of these university programs have situated a practical component in their courses, using evaluation projects to address addressing what Chelimsky (2012) identifies as a deficit of information regarding the relationship between evaluation theory and evaluation practice.

This session highlights the theoretical and practical challenges in creating safe spaces for students to explore the social, political, and cultural dynamics of evaluation. Chouinard et al. (2017) posit this student experienced tension as the ultimate debate between “instrumental value of evaluation and its technical application, or on evaluation as an expression of our sociopolitical values and beliefs about social change.” A key dimension of this debate, includes the knowledge created through practical application in project evaluation work, and how this knowledge is used by both evaluators and project clients (Schwandt, 2007). These evaluation educators share steps undertaken to build the competencies and confidence of students to ensure that they are better able to respond to the complexities required in addressing resilient society needs. Recent studies of practice are summarized by Schön “emphasize that it is characterized by contextuality, acting that is continuous with knowing, and interactivity, and that take together, these elements comprise an account of the kind of practical judgment required of an environment characterized by complexity, interdeterminacy, and the necessity to act on the situation at hand” In building capacities of emerging evaluators to respond to complexity, we like to say that we are bringing evaluation theory to life at the axiological intersection of epistemology and ontology, or – where the rubber meets the road.

The discussion will touch on the various training needs of students in such courses that vary from the methodological, to consultative skills required for effective evaluations. The discussion will also touch on the strategies for negotiating with the field settings that host student evaluation projects with respect to the role and work of evaluation students. Finally, panelists will share their perceptions regarding the necessary student supports for success.

**RT 11 At the Axiological Intersection of Epistemology and Ontology: Or, Where the Rubber Hits the Road****RC 42 - At the Axiological Intersection of Epistemology and Ontology: Or, Where the Rubber Hits the Road. The Former Student/Current Instructor Perspective**

*J. Whynot<sup>1</sup>*

<sup>1</sup> *University of Ottawa, Student, Ottawa, Canada*

This panel addresses Stream 2 of the conference theme “Rethinking evaluation methods and methodology”. The panel is comprised of a collective of evaluation academics including Deans, professors, and instructors who are consciously addressing the learning needs of the next generation of emerging evaluators. In North America, more than 48 university-based evaluation training programs are offered (LaVelle & Donaldson, 2010). A significant number of evaluators sidestep into the profession through sectors such as health, sociology and education (Stevahn, King, Ghery, & Minnema, 2005) and consequently do not hold advanced evaluation degrees and are motivated to seek advanced formal training (Christie, Quiñones, & Fierro, 2014). While these university programs are grounded in evaluation theory, a gap exists in bridging theoretical concepts to practical applications in evaluation work. Several of these university programs have situated a practical component in their courses, using evaluation projects to address addressing what Chelimsky (2012) identifies as a deficit of information regarding the relationship between evaluation theory and evaluation practice.

This session highlights the theoretical and practical challenges in creating safe spaces for students to explore the social, political, and cultural dynamics of evaluation. Chouinard et al. (2017) posit this student experienced tension as the ultimate debate between “instrumental value of evaluation and its technical application, or on evaluation as an expression of our sociopolitical values and beliefs about social change.” A key dimension of this debate, includes the knowledge created through practical application in project evaluation work, and how this knowledge is used by both evaluators and project clients (Schwandt, 2007). These evaluation educators share steps undertaken to build the competencies and confidence of students to ensure that they are better able to respond to the complexities required in addressing resilient society needs. Recent studies of practice are summarized by Schön “emphasize that it is characterized by contextuality, acting that is continuous with knowing, and interactivity, and that take together, these elements comprise an account of the kind of practical judgment required of an environment characterized by complexity, indeterminacy, and the necessity to act on the situation at hand” In building capacities of emerging evaluators to respond to complexity, we like to say that we are bringing evaluation theory to life at the axiological intersection of epistemology and ontology, or – where the rubber meets the road.

The discussion will touch on the various training needs of students in such courses that vary from the methodological, to consultative skills required for effective evaluations. The discussion will also touch on the strategies for negotiating with the field settings that host student evaluation projects with respect to the role and work of evaluation students. Finally, panelists will share their perceptions regarding the necessary student supports for success.

**RT 11 At the Axiological Intersection of Epistemology and Ontology: Or, Where the Rubber Hits the Road****RC 43 - At the Axiological Intersection of Epistemology and Ontology: Or, Where the Tubber Hits the Road. The Original Program Instructor Perspective**

*S. Montague*<sup>1</sup>

<sup>1</sup> Carleton University, School of Policy and Public Administration, Ottawa, Canada

This panel addresses Stream 2 of the conference theme “Rethinking evaluation methods and methodology”. The panel is comprised of a collective of evaluation academics including Deans, professors, and instructors who are consciously addressing the learning needs of the next generation of emerging evaluators. In North America, more than 48 university-based evaluation training programs are offered (LaVelle & Donaldson, 2010). A significant number of evaluators sidestep into the profession through sectors such as health, sociology and education (Stevahn, King, Ghery, & Minnema, 2005) and consequently do not hold advanced evaluation degrees and are motivated to seek advanced formal training (Christie, Quiñones, & Fierro, 2014). While these university programs are grounded in evaluation theory, a gap exists in bridging theoretical concepts to practical applications in evaluation work. Several of these university programs have situated a practical component in their courses, using evaluation projects to address addressing what Chelimsky (2012) identifies as a deficit of information regarding the relationship between evaluation theory and evaluation practice.

This session highlights the theoretical and practical challenges in creating safe spaces for students to explore the social, political, and cultural dynamics of evaluation. Chouinard et al. (2017) posit this student experienced tension as the ultimate debate between “instrumental value of evaluation and its technical application, or on evaluation as an expression of our sociopolitical values and beliefs about social change.” A key dimension of this debate, includes the knowledge created through practical application in project evaluation work, and how this knowledge is used by both evaluators and project clients (Schwandt, 2007). These evaluation educators share steps undertaken to build the competencies and confidence of students to ensure that they are better able to respond to the complexities required in addressing resilient society needs. Recent studies of practice are summarized by Schön “emphasize that it is characterized by contextuality, acting that is continuous with knowing, and interactivity, and that take together, these elements comprise an account of the kind of practical judgment required of an environment characterized by complexity, indeterminacy, and the necessity to act on the situation at hand” In building capacities of emerging evaluators to respond to complexity, we like to say that we are bringing evaluation theory to life at the axiological intersection of epistemology and ontology, or – where the rubber meets the road.

The discussion will touch on the various training needs of students in such courses that vary from the methodological, to consultative skills required for effective evaluations. The discussion will also touch on the strategies for negotiating with the field settings that host student evaluation projects with respect to the role and work of evaluation students. Finally, panelists will share their perceptions regarding the necessary student supports for success.

**RT 11 At the Axiological Intersection of Epistemology and Ontology: Or, Where the Rubber Hits the Road****RC 44 - At the Axiological intersection of Epistemology and Ontology: Or, Where the Rubber Hits the Road – The Management Perspective**

*J. Sylvestre*<sup>1</sup>

<sup>1</sup> *University of Ottawa, Faculty of Social Sciences, Ottawa, Canada*

This panel addresses Stream 2 of the conference theme “Rethinking evaluation methods and methodology”. The panel is comprised of a collective of evaluation academics including Deans, professors, and instructors who are consciously addressing the learning needs of the next generation of emerging evaluators. In North America, more than 48 university-based evaluation training programs are offered (LaVelle & Donaldson, 2010). A significant number of evaluators sidestep into the profession through sectors such as health, sociology and education (Stevahn, King, Ghery, & Minnema, 2005) and consequently do not hold advanced evaluation degrees and are motivated to seek advanced formal training (Christie, Quiñones, & Fierro, 2014). While these university programs are grounded in evaluation theory, a gap exists in bridging theoretical concepts to practical applications in evaluation work. Several of these university programs have situated a practical component in their courses, using evaluation projects to address addressing what Chelimsky (2012) identifies as a deficit of information regarding the relationship between evaluation theory and evaluation practice.

This session highlights the theoretical and practical challenges in creating safe spaces for students to explore the social, political, and cultural dynamics of evaluation. Chouinard et al. (2017) posit this student experienced tension as the ultimate debate between “instrumental value of evaluation and its technical application, or on evaluation as an expression of our sociopolitical values and beliefs about social change.” A key dimension of this debate, includes the knowledge created through practical application in project evaluation work, and how this knowledge is used by both evaluators and project clients (Schwandt, 2007). These evaluation educators share steps undertaken to build the competencies and confidence of students to ensure that they are better able to respond to the complexities required in addressing resilient society needs. Recent studies of practice are summarized by Schön “emphasize that it is characterized by contextuality, acting that is continuous with knowing, and interactivity, and that take together, these elements comprise an account of the kind of practical judgment required of an environment characterized by complexity, indeterminacy, and the necessity to act on the situation at hand” In building capacities of emerging evaluators to respond to complexity, we like to say that we are bringing evaluation theory to life at the axiological intersection of epistemology and ontology, or – where the rubber meets the road.

The discussion will touch on the various training needs of students in such courses that vary from the methodological, to consultative skills required for effective evaluations. The discussion will also touch on the strategies for negotiating with the field settings that host student evaluation projects with respect to the role and work of evaluation students. Finally, panelists will share their perceptions regarding the necessary student supports for success.

**RT 11 At the Axiological Intersection of Epistemology and Ontology: Or, Where the Rubber Hits the Road****RC 97 - At the Axiological Intersection of Epistemology and Ontology: Or, Where the Rubber Hits the Road. The Program Supervisor Perspective**

*J. Whynot<sup>1</sup>, S. Montague<sup>2</sup>*

<sup>1</sup> University of Ottawa, Student, Ottawa, Canada

<sup>2</sup> Carleton University, School of Policy and Public Administration, Ottawa, Canada

This panel addresses Stream 2 of the conference theme “Rethinking evaluation methods and methodology”. The panel is comprised of a collective of evaluation academics including Deans, professors, and instructors who are consciously addressing the learning needs of the next generation of emerging evaluators. In North America, more than 48 university-based evaluation training programs are offered (LaVelle & Donaldson, 2010). A significant number of evaluators sidestep into the profession through sectors such as health, sociology and education (Stevahn, King, Ghery, & Minnema, 2005) and consequently do not hold advanced evaluation degrees and are motivated to seek advanced formal training (Christie, Quiñones, & Fierro, 2014). While these university programs are grounded in evaluation theory, a gap exists in bridging theoretical concepts to practical applications in evaluation work. Several of these university programs have situated a practical component in their courses, using evaluation projects to address addressing what Chelimsky (2012) identifies as a deficit of information regarding the relationship between evaluation theory and evaluation practice.

This session highlights the theoretical and practical challenges in creating safe spaces for students to explore the social, political, and cultural dynamics of evaluation. Chouinard et al. (2017) posit this student experienced tension as the ultimate debate between “instrumental value of evaluation and its technical application, or on evaluation as an expression of our sociopolitical values and beliefs about social change.” A key dimension of this debate, includes the knowledge created through practical application in project evaluation work, and how this knowledge is used by both evaluators and project clients (Schwandt, 2007). These evaluation educators share steps undertaken to build the competencies and confidence of students to ensure that they are better able to respond to the complexities required in addressing resilient society needs. Recent studies of practice are summarized by Schön “emphasize that it is characterized by contextuality, acting that is continuous with knowing, and interactivity, and that take together, these elements comprise an account of the kind of practical judgment required of an environment characterized by complexity, indeterminacy, and the necessity to act on the situation at hand” In building capacities of emerging evaluators to respond to complexity, we like to say that we are bringing evaluation theory to life at the axiological intersection of epistemology and ontology, or – where the rubber meets the road.

The discussion will touch on the various training needs of students in such courses that vary from the methodological, to consultative skills required for effective evaluations. The discussion will also touch on the strategies for negotiating with the field settings that host student evaluation projects with respect to the role and work of evaluation students. Finally, panelists will share their perceptions regarding the necessary student supports for success.

Thursday, 4 October 2018  
11:30–13:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**RT 12 What Lasts? Exploring Sustainability: What Does It Mean? And How Can It Be Measured?**

## **What Lasts? Exploring Sustainability: What Does It Mean? And How Can It Be Measured?**

*V. Hildenwall*<sup>1</sup>

<sup>1</sup> FCG Swedish Development AB, Monitoring and Evaluation, Stockholm, Sweden

This panel will focus on: Exploring what lasts. A discussion about what sustainability means, how to achieve it, and how to measure it. We will be exploring sustainability questions generally and also how they relate to specific questions of gender equality, child rights and environment in view of climate change.

A high proportion of evaluations include sustainability questions. Indeed, one of the OECD DAC's key criteria for evaluation is sustainability. Simply put sustainability seeks to find out if the objectives of the intervention will remain once the intervention is over. However, how can this be assessed, and what does it really mean?

While sustainability may be straight forward in some thematic areas, in others what sustainability means is obscure at best, and how to measure it equally so. When can one say sustainability has been achieved? What does sustainability mean when the objective of the task is to achieve change in thinking? Is the sustainability question relevant at all? And, if relevant, how should it be defined? Further, what measurement tools are at the disposal of evaluators, who often contend with other challenges (such as short and single field missions, etc)? Our exploration of the issue of sustainability has led to single out a number of fields where answering the aforementioned questions is particularly difficult. These include gender equality, child rights, climate change adaptation and refugee/returnee reintegration.

This panel would like to dissect the question of sustainability and discuss it by exploring it from multiple angles: Gender equality, Child rights, Migration/refugees and Climate change/environment. This varied set of thematic areas will drive a rich and pragmatic – rather than theoretical – discussion about the meaning and implications of sustainability for evaluators, how to conceptualize change when it requires a fundamental review of attitudes, practices and approaches, and how to provide evaluations' users with rich, but realistic, insights regarding what the future might hold.

We have chosen a round table approach because we feel that the insights that each panellist brings can lead to a rich and open discussion, that includes the audience and where we can, in a collaborative and participatory manner.

**RT 12 What Lasts? Exploring Sustainability: What Does It Mean? And How Can It Be Measured?****RC 45 - Critical Times and Critical Questions: What Does Sustainability Mean for Climate Change Focused Interventions**

A. Millard<sup>1</sup>

<sup>1</sup> Policy Research Institute, Consultant, Chamoson, Switzerland

Climate change is a very current and important issue. However, in the midst of the political discussions, it is often difficult to stay true to the importance of achieving sustainable change. What does sustainability mean for climate change efforts: maintaining what has been achieved? Expanding from what has been achieved? Ensuring that there is a fundamental shift in how environmental interventions are addressed? In a world where the level of change that humanity needs to reach to curtail the impact of climatic variations, what role do evaluators play in helping evaluated parties and clients(donors) determine if what has been done will be enough in the long term? Her discussion points and intervention will aim to identify some key issues from current evaluations and ask critical questions about how evaluators may better prepare to tackle sustainability questions in this sector.

**RT 12 What Lasts? Exploring Sustainability: What Does It Mean? And How Can It Be Measured?****RC 46 - Evaluating Sustainability in Relation to Children's Rights**

T. Ling<sup>1</sup>

<sup>1</sup> RAND corporation, Evaluation, Cambridge, United Kingdom

The language and philosophy of child rights in development asserts two things; children are viewed as rights holders and others are identified as duty bearers whose duty it is to provide these rights. This introduced a language emphasising the moral necessity of development rather than charity and helped change the discourse of development in arguably very helpful ways. However, in a short presentation, I wish to explore the possibility that a rights-based assertion of immutable rights and obligations, has created a linguistic gloss around development activity that makes it harder to evaluate sustainability which aims to measure whether the benefits of an activity are likely to continue after donor funding has been withdrawn (including both environmental as well as financial sustainability). Children's rights and duty bearer's obligations can be hindered or unfulfilled for any number of reasons which are unrelated to the programme being evaluated. Either a capability approach or a welfare approach might focus on the sustainability of certain capabilities or benefits and these might be measured. In practice this is what we choose to measure in our evaluations. Is this a problem for a rights based approach?

**RT 12 What Lasts? Exploring Sustainability: What Does It Mean? And How Can It Be Measured?****RC 47 - Children Rights and Gender: Where the Two Roads Meet**

S. Rauh<sup>1</sup>

<sup>1</sup> Policy Research Institute, Consultant, Chamoson, Switzerland

Child rights and gender are often treated as two separate thematic areas. However, it is clear that creating social change can often be a generational question. Therefore exploring gender dynamics and supporting gender equality amongst children is key. It is widely documented that gendered perspectives and prejudices are developed amongst the young who mimic the attitudes and practices of the adults around them. This means that interventions need to account for negative reinforcement and that a sustainable effort can only be truly seen many years later. Is it relevant to measure sustainability at the end of a project or programme? What type of paths need to be forged to ensure that efforts will materialize as they are expected. These will be the core questions that my short exposition will put forward.

**RT 12 What Lasts? Exploring Sustainability: What Does It Mean? And How Can It Be Measured?****RC 48 - Solutions and Sustainable Reintegration: What Are We After and How Do We Identify It?**

Y. Hasegawa<sup>1</sup>

<sup>1</sup> Policy Research Institute, Consultant, Chamoson, Switzerland

Refugees and returnees and the dynamics around the issue are complex and often overshadowed by the “emergency” aspect of the intervention. However, ensuring that efforts are sustainable and that reintegration does actually happen is central. It’s important to know how sustainability can be achieved so that lessons learned can be replicated. But this means we must measure sustainability. What does sustainability mean in a post-crisis environment? This short presentation will explore the challenges that the subject matter poses and set out a number of questions regarding how interventions are designed and the degree to which sufficient focus is placed on exploring sustainability. These concepts will guide the 10–15 min presentation I will give to frame the discussion.

**RT 12 What Lasts? Exploring Sustainability: What Does It Mean? And How Can It Be Measured?****RC 49 - Sustainability of Gender Equality Interventions**

S. Callegari<sup>1</sup>

<sup>1</sup> United Nations, United Nations Department of Safety and Security, New York, USA

Gender equality programming by governmental and non-governmental actors alike has experienced several “silver bullet” phases in which practitioners had felt that they had finally identified the definite and sustainable solution to the empowerment of women and gender equality. We all remember the “microcredit” phase, the “basic education” phase, etc. And while those magic bullets have left a mark and, in most cases, made a positive difference in the life of the women and girls (but also men and boys) they touched, sustainability has always seemed elusive, in particular under the short-term spectrum of evaluations happening right at the end of a programme. The “silver bullets” phases have taught a lot in terms of what has impact, how to find it and how to measure it. The presentation will make reference to these elements, based on lessons learnt and findings both from evaluative and implementation work.

Thursday, 4 October 2018  
11:30–13:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

### RT 13 Accountability in Evaluation for Sustained and Resilient Impacts

## Accountability in Evaluation for Sustained and Resilient Impacts

*S. Bayley*<sup>1</sup>

<sup>1</sup> *Department of Foreign Affairs and Trade, Performance Management and Results, Canberra, Australia*

Accountability is "an obligation or willingness to accept responsibility or to account for one's actions" and resilience requires "the ability of an ecosystem to return to its original state after being disturbed". In international development, resilience against shocks presumes that our interventions have been beneficial while we intervened and sustained after we left. Donors are typically accountable to taxpayers and investors, governments to their citizens and NGOs to both their funders and to their participants. How well are we doing? We suggest that despite continuous evaluations over the decades, the same or almost similar interventions are being repeated around the world with limited effect, which is evident from very slow pace of progress and little evidence of sustained impacts in different sectors. This seems to point to a lack of sufficiently good design and implementation, funding and political will around what is needed for sustained and resilient impacts, especially in the face of new shocks, including climatic, economic and political ones. What needs to change? We posit that evaluations are key feedback mechanisms to show accountability to our ultimate clients, those we aim to serve in our global development projects. There could be two reasons for that – one, the organizations are not making proper use of the evaluations; two, the evaluations are failing to take account of those important issues (Datta, 2006). There is little discussion about the accountability of evaluation and evaluators to participants.

Two roundtable presenters will present and then the discussant will discuss:

- 1) Scott Bayley: Chair/ Presenter: What are the challenges evaluation faces from funder's perspectives hoping to foster sustained impacts in this resilience-challenged world? (15 mins)

*Clarifying questions: up to 5 minutes*

- 2) Jindra Cekan: Presenter: What hinders global development from evaluating sustained impacts and being accountable to our participants and partners? What can be unblocked so that partnerships that really further resilient 'development' emerge? (15 mins)

*Clarifying questions: up to 5 minutes*

- 3) Burt Perrin: Discussant: Accountability to whom? By whom? And for what? Contrasting the top down donor/funder model of accountability to them, with the civil society model of accountability to the community (15 mins)

*Clarifying questions: up to 5 minutes*

- 4) ALL Participants: We are designing the discussion to be iterative, and will lead to better evaluations, better design, implementation, monitoring and funding for more sustainable and resilient communities. *At the end of each short presentation, clarifying questions will be addressed, but the focus of the roundtable will be the discussion by participants in the room on issues brought up such as other challenges public/ private funders face, what hinders sustained impacts from being accountable to country nationals and how to broker a discussion between two seemingly opposing views. We will have 30 minutes for discussion and brainstorming what to change in their organizations or in their own work, in order to foster change.*

**RT 13 Accountability in Evaluation for Sustained and Resilient Impacts****RC 50 - Accountability in Evaluation for Sustained and Resilient Impacts: the Funders' View**

S. Bayley<sup>1</sup>

<sup>1</sup> Department of Foreign Affairs and Trade, Performance Management and Results, Canberra, Australia

INDIVIDUAL PANEL PRESENTATION on Accountability in Evaluation for Sustained and Resilient Impacts: the Funders' View What are the challenges evaluation faces from funder's perspectives hoping to foster sustained impacts in this resilience-challenged world? There is a great need to change the conversation amongst influential evaluators, get them talking about the importance of sustainability over the long-term. This presentation will discuss the barriers to truly 'sustainable development' that public funders face, yet also will open the space for discussion on how we collectively can lead to practice changes in international development organisations.

**RT 13 Accountability in Evaluation for Sustained and Resilient Impacts****RC 51 - Accountability in Evaluation for Sustained and Resilient Impacts: A View from NGOs and Participants**

J. Cekan/ova<sup>1</sup>

<sup>1</sup> Valuing Voices at Cekan Consulting LLC, President/ Catalyst, Prague, Czechia

Jindra Cekan's Presentation comes from 30 years of working with NGOs and listening to participants. We often ignore accountability, as evidenced by far fewer than 1% of all projects being evaluated for sustained (much less emerging) impacts after close-out. How much more resilient can our participants be if our projects have not strengthened their ability to thrive even a few years after our projects end. How many sectors learn from their evaluations and fundamentally change programming? We mostly lack political will around what is needed to foster sustained and resilient impacts with our participants, especially in the face of new shocks, including climatic, economic and political ones. What hinders global development from evaluating sustained impacts and being accountable to our participants and partners? What can be unblocked so that partnerships that really further resilient 'development' emerge?

**RT 13 Accountability in Evaluation for Sustained and Resilient Impacts****RC 52 - Round-table Contribution**

B. Perrin<sup>1</sup>

<sup>1</sup> Independent Consultant, Vissec, France

Burt Perrin, Burt will act as discussant, considering multiple meanings of accountability and raising questions such as: "accountability to whom? by whom? and for what?". He will also add in some considerations from his recent paper about the two competing means of accountability by donors and civil society, and will pose some questions for the roundtable discussion to follow.

Thursday, 4 October 2018  
15:00–16:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

### RT 15 New Frontiers in Development Evaluation: the case of Blended Finance

## New Frontiers in Development Evaluation: The Case of Blended Finance

*I. Basile*<sup>1</sup>

<sup>1</sup> OECD, Development Co-operation Directorate, Policy Analyst, Paris, France

Private finance plays an increasingly important role in bringing innovation, expertise and additional resources to help developing countries achieve the Sustainable Development Goals. Recognising this, development co-operation providers are increasingly working with the private sector to unlock commercial finance. As Blended Finance becomes an integral part of donor countries' development cooperation strategies, development finance institutions and other implementing agencies will have to ensure similar levels of accountability, especially to policy makers and in the public eye.

The 2018 OECD report "Making Blended Finance Work for the Sustainable Development Goals" presents several findings relevant to the monitoring and evaluation community:

- (i) National evaluation strategies on development cooperation fail to address the complex governance and high intermediation of blended finance. Blended Finance practitioners lack of a common vocabulary and understanding concerning development results.
- (ii) The high fragmentation in the governance of blended finance presents a challenge for monitoring and evaluation: over 20% of the blended finance funds and facilities surveyed in 2017 have no formalised internal M&E function.
- (iii) Evaluation practice varies greatly depending on the vehicle, with 81% of the surveyed blended finance facilities undertaken at least one evaluation compared to 56% of funds. The causal link is rarely tested, thereby limiting the learning potential for decision makers. Less than half of final evaluation reports from surveyed facilities and funds are made public.

The roundtable aims to: (1) Foster the exchange of information and experiences among evaluation practitioners, (2) Identify the strengths and weaknesses of different evaluation techniques when applied to Blended Finance, (3) Discuss which evaluation techniques are most appropriate depending on the instruments and governance of diverse Blended Finance mechanisms.

Each presentation shall examine the strengths and weaknesses of a different evaluation methodologies applied to Blended Finance in practice. Discussants will reflect upon the robustness of their findings on development effectiveness and impacts, the coverage of other evaluation criteria (incl. evidence on additionality, crowding out and value for money), the ownership of evaluation conclusions by blended finance managers and investors, the learning potential for the evaluation profession.

The international community shall benefit from adopting a learning agenda on blended finance to guide the development of monitoring and evaluation practice. By adopting the OECD Blended Finance Principles for Unlocking Commercial Finance for the SDGs, members of the Development Assistance Committee have committed to monitoring development results of their blended finance operations. This discussion is part of a wider consultation process, whereby the OECD Development Cooperation Directorate will establish further policy guidance on each Blended Finance Principle.

**RT 15 New Frontiers in Development Evaluation: the case of Blended Finance**

**RC 53 - The OECD DAC Blended Finance Principles for Unlocking Commercial Finance for the SDGs**

I. Basile<sup>1</sup>

<sup>1</sup> OECD, Development Co-operation Directorate, Paris, France

Irene Basile will present key findings from the 2018 OECD report “Making Blended Finance Work for the Sustainable Development Goals” and the ongoing consultation process around the OECD DAC Blended finance principles for unlocking commercial finance for the SDGs.

**RT 15 New Frontiers in Development Evaluation: the case of Blended Finance**

**RC 54 - Methodological Issues on the Evaluation of Blended Finance**

O. Winckler Andersen<sup>1</sup>

<sup>1</sup> Danish Institute for International Studies, Senior Analyst, Copenhagen, Denmark

Ole Winckler Andersen will discuss the existing capacity gaps on blended finance evaluation, and the urgency to establish a shared evaluation framework to ensure comparability with other development aid modalities.

**RT 15 New Frontiers in Development Evaluation: the case of Blended Finance**

**RC 55 - Evaluating the Use of Subsidies in a Development Bank**

O. Mrinska<sup>1</sup>

<sup>1</sup> EBRD, Principal Evaluation Manager, London, United Kingdom

Olga Mrinska will share the experience in conducting a number of the theory-based evaluations, including at the facility level, where EBRD’s loans are blended with the investment grants, concessional loans, financing for incentives and risk sharing provided by various donors.

**RT 15 New Frontiers in Development Evaluation: the case of Blended Finance**

**RC 56 - Independent Evaluation of EU Blending Facilities (2007–2013)**

E. Buhl-Nielsen<sup>1</sup>

<sup>1</sup> PEM Consult, Partner, Düsseldorf, Germany

<sup>5</sup> Steward Redqueen, Consultant, Amsterdam, Netherlands

Eric Buhl-Nielsen will relate his experience leading the external theory-based evaluation of EU support through blending from 2007 to 2013 for the European Commission.

**RT 15 New Frontiers in Development Evaluation: the case of Blended Finance**

## **RC 57 - Insights and Lessons Learned from Evaluations of Impact Investment Funds**

*M. de Bruijn*<sup>1</sup>

<sup>1</sup> *Steward Redqueen, Director, Amsterdam, Netherlands*

Matthijs de Bruijn will share insights from evaluations of impact investment funds, drawing on his experience with several development finance institutions (Proparco, CDC Group, DEG) and exploring the issue of attribution in impact investment.

Thursday, 4 October 2018  
15:00–16:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**RT 16 Meaningful Monitoring and Evaluation systems through Evidence-Based Leadership and Evolutionary Purpose**

## Meaningful Monitoring and Evaluation Systems Through Evidence-Based Leadership and Evolutionary Purpose

A. Guidoccio<sup>1</sup>, V. Smith<sup>2</sup>

<sup>1</sup> Ministry of Modernization- National Government Argentina, Undersecretary of Open Government and Public Innovation, Buenos Aires, Argentina

<sup>2</sup> data2insight, Founder, Seattle, USA

**Justification:** In this age of increasing complexity, rapid change, and demand for greater transparency and accountability from leaders the most successful teams of the future will excel at practical performance measurement and self-management. We are eager to share two frameworks with participants so they can use them to transform their mindset and practices, clarify and measure their organization's mission, vision, and purpose, and use credible, actionable evidence to continually improve.

**Objectives:** One of the biggest challenges of putting in place a successful monitoring and evaluation (M&E) system is designing and defining meaningful measures that people use to improve program and team performance. This special activity shares concepts and tools to help evaluators and stakeholders:

to assess where their organizations are now in terms of producing and using meaningful evidence for improving performance; and

define the actions and changes needed to create learning organizations where teams are continually improving and achieving the results they truly desire.

Participants will leave with resources about, and a better understanding of evidence based leadership, self-organization and self-management as well as how they can use these tools to become high performing teams.

It is common to see leaders and M&E units making the mistake of introducing practices (e.g. annual planning, KPI formulation processes, monthly reporting, and technological tools like performance dashboard) that soon are perceived as useless or complicated routines that do not provide useful performance measurement. Specifically, we have seen problems like teams and leaders who are unclear about their purpose, priorities and the results they want to achieve. Or leaders receiving misleading evidence and no actionable recommendations from KPIs. These problems often result in teams not using the data received through monitoring systems or not sharing data adequately. What a waste of time and money!

**Rationale:** How to get out of this trap? We draw from our team's experience leveraging Performance Specialist Stacey Barr's six habits of evidence-based leaders and Frederic Laloux's concepts in his book "Reinventing organizations" to show participants how to develop and implement a monitoring system that provides timely, relevant, credible, and actionable evidence that stakeholders could use to continually improve.

Session outline

Introduce the six habits of evidence-based leadership

Describe the framework of "Teal organizations" to improve performance and collaboration through "self-management, wholeness, and evolutionary purpose."

Review two assessment instruments evaluators can use with their own teams and/or with clients:

1. PuMP Diagnostic Tool for assessing how your organization's performance measurement practices are on these five dimensions: measurable strategy, meaningful measures, measure implementation, meaningful measures, and improve performance;

2. Teal Organization Readiness Assessment for assessing your organization's processes, strategy, and culture organized around four operational categories: structure, human resources, daily life, and organizational practice;

Review the two steps for designing and defining measures in a participatory way to make them useful to staff and stakeholders; and

Discuss best practices for teams to make high-quality evidence-based decisions and taking action that results in better outcomes.

**RT 16 Meaningful Monitoring and Evaluation systems through Evidence-Based Leadership and Evolutionary Purpose**

**RC 58 - Part 1: How to Build a Career in the Evaluation Field**

A. Guidoccio<sup>1</sup>

<sup>1</sup> Ministry of Modernization- National Government Argentina, Undersecretary of Open Government and Public Innovation, Buenos Aires, Argentina

Explanation of the main competencies of evaluators and the common challenges and skills needed to acquire them.

**RT 16 Meaningful Monitoring and Evaluation systems through Evidence-Based Leadership and Evolutionary Purpose**

**RC 59 - Part 2: Self- Assessment of Evaluators Competences**

A. Guidoccio<sup>1</sup>

<sup>1</sup> University of Connecticut, Assistant Professor of Measurement- Evaluation- and Assessment, Connecticut, USA

To align with the European context, we will explain the Capabilities Framework developed by UKES and adapted by EES, and how this framework can be used for self-assessment and reflective practice. In doing so, we will share our real-life experience adapting the "Essential Competencies for Program Evaluators Self- Assessment" Stevahn, King, Ghere, and Minnema (2004) for EvalYouth's Mentoring Program. Next, we will divide the audience into groups, provide hand-outs of the EES self-assessment, and ask them to fill it individually. Small group processing and whole group share-out will be facilitated and enhanced through the use of an appreciative inquiry process facilitated by the chairs and other contributors.

**RT 16 Meaningful Monitoring and Evaluation systems through Evidence-Based Leadership and Evolutionary Purpose**

**RC 60 - Part 3: Formulation of a Personal Development Plan**

M. Branco<sup>1</sup>

<sup>1</sup> Economist and Evaluation Specialist, City, USA

Discussion on the importance of defining a career path with meaningful and specific targets. We will provide a template for a professional plan, explain its components and ask the participants to fill it individually. YEEs could prioritize which ones are their aspirations, strengths and weaknesses; and get a better understanding of what she/he needs to do to get there.

**RT 16 Meaningful Monitoring and Evaluation systems through Evidence-Based Leadership and Evolutionary Purpose**

**RC 61 - Part 4: Useful resources**

*A. Rishko-Porcescu<sup>1</sup>*

<sup>1</sup> *an independent M&E specialist in social sphere, USA*

Discussion of useful resources, opportunities, and steps to find opportunities in their area of interest.

Thursday, 4 October 2018  
17:30–19:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**RT 17 Evaluation Capacity Development: a Reflection on the Past, Present and Future of the International Program for Development Evaluation Training (IPDET)**

## **Evaluation Capacity Development: a Reflection on the Past, Present and Future of the International Program for Development Evaluation Training (IPDET)**

*J. Vaessen*<sup>1</sup>

<sup>1</sup> World Bank, Independent Evaluation Group, Washington DC, USA

The world around the practice of evaluation is rapidly evolving, and evaluation has to change with it. Within an increasingly complex and interconnected world, with growing but very uneven levels of institutionalization of evaluation across organizations, sectors, and countries, with new opportunities and challenges arising from emerging technologies and data, with profound changes in the role of external development assistance, with growing capacities in organizations in the Global South, and so on, the demands and requirements for evaluation in the field of international development are changing. Evaluation capacity development programs need to adapt in order to be relevant and effective in training qualified evaluation practitioners. The International Program for Development Evaluation Training (IPDET) is one of the longest standing and most successful evaluation training programs in the field of international development. Founded in 2001, with more than 2600 alumni from 161 countries, IPDET has made an important contribution to evaluation capacities in the world. Yet, since then the world has changed, evaluation has changed and the world of evaluation capacity development has changed, with more and more training programs on evaluation being delivered all over the world. The purpose of this round-table is to discuss the evolving practice of evaluation capacity development and to draw lessons from IPDET's rich experience as it moves into a new phase of evaluation capacity development. The round-table brings together some of the key actors of IPDET's past, present and future. At the core of the round-table is a discussion of the following questions: 1 What have been some of the key factors that made IPDET successful? What have been some of the key challenges and constraints? What can we learn from these? 2 What are the key factors that have a bearing on IPDET's future role in evaluation capacity development (e.g. relating to changes in the field of international development and the global community in general, changes in evaluation as an evolving field of practices, the growing number of evaluation trainings across the globe, etc.)? 3 Finally, looking forward, how can IPDET be as relevant, effective and inclusive as possible in meeting the training needs of evaluation stakeholders? What changes in the curriculum, teaching methods, delivery modes and locations, and/or the involvement of evaluation stakeholders from across the globe, are needed? The panel is structured as follows: After an introduction by the chair, one of the panel members will be invited to present a lightning talk (5–7 minutes) on question 1. The chair will then ask follow-up questions to the panel member and will also request concise feedback from the other panel members. The chair will then invite a second panel member to introduce question 2 and the process will be repeated. This is followed by a third panel member who will talk about question 3 and again the process is repeated. This will take up an estimated two thirds (60 minutes) of the session. In the remainder of the session the debate will be opened up for questions and answers with the audience.

**RT 17 Evaluation Capacity Development: a Reflection on the Past, Present and Future of the International Program for Development Evaluation Training (IPDET)**

## **RC 62 - Round-table Contribution**

R. Rist<sup>1</sup>

<sup>1</sup> Independent Consultant, International Program for Development Evaluation Training, Washington DC, USA

**RT 17 Evaluation Capacity Development: a Reflection on the Past, Present and Future of the International Program for Development Evaluation Training (IPDET)**

## **RC 63 - Round-table Contribution**

W. Meyer<sup>1</sup>

<sup>1</sup> University of Saarland, Ceval, Saarbrücken, Germany

**RT 17 Evaluation Capacity Development: a Reflection on the Past, Present and Future of the International Program for Development Evaluation Training (IPDET)**

## **RC 64 - Round-table Contribution**

J. Vaessen<sup>1</sup>

<sup>1</sup> World Bank, Independent Evaluation Group, Washington DC, USA

Thursday, 4 October 2018  
17:30–19:00

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**RT 18 How to build a career in evaluation?: Assess your competencies and develop a professional development plan. Roundtable or unconference event**

## How to Build a Career in Evaluation?: Assess Your Competencies and Develop a Professional Development Plan. Roundtable or Unconference Event

*A. Guidoccio*<sup>1</sup>

<sup>1</sup> Ministry of Modernization- National Government Argentina, Undersecretary of Open Government and Public Innovation, Buenos Aires, Argentina

**Rationale:** One of the priorities of the Evaluation Agenda 2020 is to build capacities for evaluation. This is particularly important in countries where the production and use of evaluation are limited (e.g., Argentina) and there are few options for evaluation training (e.g., Finland). This activity proposes an innovative way to contribute to this objective with a focus on the individual capacities of novice evaluators.

**Objectives:** Based on the work of EvalYouth and the materials of its International Mentoring Program, this activity targets Young and Emerging Evaluators (YEEs) and aims to help them to assess their competencies as evaluators; clarify their specific interest within the field; develop a professional plan to guide their future career paths; and acquire new resources to succeed in the evaluation career.

Through a participatory approach and work in groups, it also aims to empower the participants with new connections and exchange of experiences with other YEEs and invite them to be part of a community of learning.

The activity will be divided into 4 parts. Facilitators are leaders of EvalYouth Global, two EvalYouth regional chapters, and the EES Emerging Evaluators TWG.

Part 1: How to build a career in the evaluation field?

Explanation of the main competencies of evaluators and the common challenges and skills needed to acquire them.

Part 2: Self- Assessment of evaluators competences. To align with the European context, we will explain the Capabilities Framework developed by UKES and adapted by EES, and how this framework can be used for self-assessment and reflective practice. In doing so, we will share our real-life experience adapting the “Essential Competencies for Program Evaluators Self- Assessment” Stevahn, King, Ghere, and Minnema (2004) for EvalYouth’s Mentoring Program. Next, we will divide the audience into groups, provide handouts of the EES self-assessment, and ask them to fill it individually. Small group processing and whole group share-out will be facilitated and enhanced through the use of an appreciative inquiry process facilitated by the chairs and other contributors.

Part 3: Formulation of a Personal Development Plan

Discussion on the importance of defining a career path with meaningful and specific targets. We will provide a template for a professional plan, explain its components and ask the participants to fill it individually. YEEs could prioritize which ones are their aspirations, strengths and weaknesses; and get a better understanding of what she/he needs to do to get there.

Part 4: Useful resources

Discussion of useful resources, opportunities, and steps to find opportunities in their area of interest. Justification: EvalYouth conducted a needs assessment in 2016 with over 300 respondents from 69 countries. Over 90% described mentoring as one of the most important unmet needs. EvalYouth has developed several strategies in response to this need, including its’ mentoring program.

It also presents an opportunity for formal collaboration and stronger ties across several groups working on the needs of novice evaluators (e.g., EvalYouth, EES TWG 5). This is in recognition that we are all working toward the same end, and our collective work is stronger together.

**RT 18 How to build a career in evaluation?: Assess your competencies and develop a professional development plan. Roundtable or unconference event**

## **RC 65 - How to Build a Career in Evaluation?: Assess Your Competencies and Develop a Professional Development Plan. Roundtable or Unconference Event**

*A. Guidoccio<sup>1</sup>, B. Montrosse-Moorhead<sup>2</sup>, M. Branco<sup>3</sup>, A. Rishko<sup>4</sup>*

<sup>1</sup> Ministry of Modernization- National Government Argentina, Undersecretary of Open Government and Public Innovation, Buenos Aires, Argentina

<sup>2</sup> University of Connecticut, Department of Educational Psychology, Connecticut, USA

<sup>3</sup> Mariana Branco Consultant, Consultant, Portugal, Portugal

<sup>4</sup> Consultant, Consultant, Moldova, Moldova Republic of

**Rationale:** One of the priorities of the Evaluation Agenda 2020 is to build capacities for evaluation. This is particularly important in countries where the production and use of evaluation are limited (e.g., Argentina) and there are few options for evaluation training (e.g., Finland). This activity proposes an innovative way to contribute to this objective with a focus on the individual capacities of novice evaluators.

**Objectives:** Based on the work of EvalYouth and the materials of its International Mentoring Program, this activity targets Young and Emerging Evaluators (YEEs) and aims to help them to assess their competencies as evaluators; clarify their specific interest within the field; develop a professional plan to guide their future career paths; and acquire new resources to succeed in the evaluation career.

Through a participatory approach and work in groups, it also aims to empower the participants with new connections and exchange of experiences with other YEEs and invite them to be part of a community of learning.

The activity will be divided into 4 parts. Facilitators are leaders of EvalYouth Global, two EvalYouth regional chapters, and the EES Emerging Evaluators TWG.

Part 1: How to build a career in the evaluation field?

Explanation of the main competencies of evaluators and the common challenges and skills needed to acquire them.

Part 2: Self- Assessment of evaluators competences

To align with the European context, we will explain the Capabilities Framework developed by UKES and adapted by EES, and how this framework can be used for self-assessment and reflective practice. In doing so, we will share our real-life experience adapting the "Essential Competencies for Program Evaluators Self- Assessment" Stevahn, King, Ghere, and Minnema (2004) for EvalYouth's Mentoring Program. Next, we will divide the audience into groups, provide hand-outs of the EES self-assessment, and ask them to fill it individually. Small group processing and whole group share-out will be facilitated and enhanced through the use of an appreciative inquiry process facilitated by the chairs and other contributors.

Part 3: Formulation of a Personal Development Plan

Discussion on the importance of defining a career path with meaningful and specific targets. We will provide a template for a professional plan, explain its components and ask the participants to fill it individually. YEEs could prioritize which ones are their aspirations, strengths and weaknesses; and get a better understanding of what she/he needs to do to get there.

Part 4: Useful resources

Discussion of useful resources, opportunities, and steps to find opportunities in their area of interest.

Justification: EvalYouth conducted a needs assessment in 2016 with over 300 respondents from 69 countries. Over 90% described mentoring as one of the most important unmet needs. EvalYouth has developed several strategies in response to this need, including its' mentoring program.

It also presents an opportunity for formal collaboration and stronger ties across several groups working on the needs of novice evaluators (e.g., EvalYouth, EES TWG 5). This is in recognition that we are all working toward the same end, and our collective work is stronger together.

**RT 18 How to build a career in evaluation?: Assess your competencies and develop a professional development plan. Roundtable or unconference event**

## **RC 66 - How to Build a Career in Evaluation?: Assess Your Competencies and Develop a Professional Development Plan. Roundtable or Unconference Event**

A. Guidoccio<sup>1</sup>, B. Montrosse-Moorhead<sup>2</sup>, M. Branco<sup>3</sup>, A. Rishko<sup>4</sup>

<sup>1</sup> Ministry of Modernization- National Government Argentina, Undersecretary of Open Government and Public Innovation, Buenos Aires, Argentina

<sup>2</sup> University of Connecticut, Department of Educational Psychology, Connecticut, USA

<sup>3</sup> Mariana Branco Consultant, Consultant, Portugal, Portugal

<sup>4</sup> Consultant, Consultant, Moldova, Moldova Republic of

**Rationale:** One of the priorities of the Evaluation Agenda 2020 is to build capacities for evaluation. This is particularly important in countries where the production and use of evaluation are limited (e.g., Argentina) and there are few options for evaluation training (e.g., Finland). This activity proposes an innovative way to contribute to this objective with a focus on the individual capacities of novice evaluators.

**Objectives:** Based on the work of EvalYouth and the materials of its International Mentoring Program, this activity targets Young and Emerging Evaluators (YEEs) and aims to help them to assess their competencies as evaluators; clarify their specific interest within the field; develop a professional plan to guide their future career paths; and acquire new resources to succeed in the evaluation career.

Through a participatory approach and work in groups, it also aims to empower the participants with new connections and exchange of experiences with other YEEs and invite them to be part of a community of learning.

The activity will be divided into 4 parts. Facilitators are leaders of EvalYouth Global, two EvalYouth regional chapters, and the EES Emerging Evaluators TWG.

Part 1: How to build a career in the evaluation field?

Explanation of the main competencies of evaluators and the common challenges and skills needed to acquire them.

Part 2: Self- Assessment of evaluators competences

To align with the European context, we will explain the Capabilities Framework developed by UKES and adapted by EES, and how this framework can be used for self-assessment and reflective practice. In doing so, we will share our real-life experience adapting the "Essential Competencies for Program Evaluators Self- Assessment" Stevahn, King, Ghere, and Minnema (2004) for EvalYouth's Mentoring Program. Next, we will divide the audience into groups, provide hand-outs of the EES self-assessment, and ask them to fill it individually. Small group processing and whole group share-out will be facilitated and enhanced through the use of an appreciative inquiry process facilitated by the chairs and other contributors.

Part 3: Formulation of a Personal Development Plan

Discussion on the importance of defining a career path with meaningful and specific targets. We will provide a template for a professional plan, explain its components and ask the participants to fill it individually. YEEs could prioritize which ones are their aspirations, strengths and weaknesses; and get a better understanding of what she/he needs to do to get there.

#### Part 4: Useful resources

Discussion of useful resources, opportunities, and steps to find opportunities in their area of interest.

Justification: EvalYouth conducted a needs assessment in 2016 with over 300 respondents from 69 countries. Over 90% described mentoring as one of the most important unmet needs. EvalYouth has developed several strategies in response to this need, including its' mentoring program.

It also presents an opportunity for formal collaboration and stronger ties across several groups working on the needs of novice evaluators (e.g., EvalYouth, EES TWG 5). This is in recognition that we are all working toward the same end, and our collective work is stronger together.

Thursday, 4 October 2018  
17:30–19:00

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**RT 19 The System for Evaluating International Development Cooperation in Sweden: Is the System Fit for Purpose?**

## The System for Evaluating International Development Cooperation in Sweden: Is the System Fit for Purpose?

*V. Hildenwall<sup>1</sup>*

<sup>1</sup> FCG Sweden, M&E Unit, Stockholm, Sweden

Delivering intended change through public policy is challenging in any context, but particularly so in international development cooperation, where evaluation is becoming more and more important, particular in these turbulent times. Yet, demonstrating and learning from evidence around what works in the real world is extremely difficult. In order to deal with these issues, effective evaluation systems for international development cooperation are required. For this round-table, we propose to discuss different aspects of effective evaluation systems, from the perspective of the Swedish system for evaluating international development cooperation, and the extent to which this system is fit for purpose. The round-table will draw on a recent OECD/DAC study on evaluation systems (2016) that states that “an effective evaluation system is one in which evaluation findings contribute to the design of programmes and projects, to policy making, and to overall organisation learning and capacity building. An effective system depends mainly on: the credibility and use of evaluation findings; whether learning is taking place; whether evaluation products are of high quality; support by senior management with implementation and follow-up of recommendations; and whether evaluation findings are made public.” The round-table would base its discussion on developments in Sweden since a 2012 review by the Swedish Agency for Public Management, which led to a broader discussion about the role of evaluation in international development cooperation in Sweden, and the subsequent replacement of the Swedish Agency for Development Evaluation (SADEV), with the new the Expert Group for Aid Studies (EBA). The report in 2012, mirroring criticism from recent OECD/DAC peer reviews, concluded that the evaluation system as a whole was not able to perform the basic requirements of an evaluation system, in this case deemed to be to support regular results based management and to give independent assessments on the effectiveness and efficiency of ODA. It was also concluded that SADEV was un-fit for purpose: the long-term impact of development cooperation had not been evaluated; there was a lack of relevance and quality of evaluations, leading to a lack of evidence for better policy making. In addition, Sida's mandate for evaluation was deemed to be unclear, specifically the role of more strategic evaluation, with the focus of activities largely internal. The objective of the roundtable would be to share Swedish experiences of broader interest to the evaluation community. This will include discussing the previous challenges in the Swedish evaluation system, how they were dealt and whether the system has improved in the last few years as a result. In addition, the round-table will also discuss what Sweden can learn from other evaluation system configurations, in this instance Germany. The round-table will be of interest to evaluation commissioners, providers and users within international development cooperation from Europe and internationally, in terms of discussing how to best configure evaluations systems in order to support greater development effectiveness, and thus advancing the public interest by ensuring better use of taxpayers money.

**RT 19 The System for Evaluating International Development Cooperation in Sweden: Is the System Fit for Purpose?****RC 67 - The System for Evaluating International Development Cooperation in Sweden: Is the System Fit for Purpose?**

*V. Hildenwall*<sup>1</sup>

<sup>1</sup> FCG Sweden, M&E Unit, Stockholm, Sweden

As chair, Viktoria will introduce the round-table participants and provide an introduction to the topic in the form of a short historical overview of the development of the evaluation system in Sweden, and the recent debates since 2012, with the replacement of SADEV with EBA and other developments. As a previous employee of SADEV and the manager of a framework contract for providing project and programme evaluations for Sida, she is in a good position to provide this overview. She will speak for 10 minutes and then pass the word to one of the discussants. After all the discussants have presented their initial contributions of 10 minutes each, she will moderate a discussion around the following questions, based on the 2016 OECD/DAC review definition of an effective evaluation system:

1. To what extent do the evaluation findings contribute to the design of programmes and projects and policy making in Sweden, and in Germany?
2. To what extent do the evaluation findings contribute to overall organisation learning and capacity building in Sweden, and in Germany?
3. To what extent is there support by senior management with implementation and follow-up of evaluation recommendations in Sweden, and in Germany?
4. Are evaluations products of high quality and of relevance?
5. To what extent are evaluation findings made public in Sweden, and in Germany?
6. Can we say that the Swedish system as a whole is better able to provide the Swedish government and its aid institutions with the accountability and learning required than in 2012? In short, does the system contribute to greater development effectiveness?
7. What can Sweden learn from the evaluation system in Germany and the way in which DEval works?
8. Are there different and more innovative ways in which the system could be organised to support better outcomes?

**RT 19 The System for Evaluating International Development Cooperation in Sweden: Is the System Fit for Purpose?****RC 68 - The Perspective of the Swedish International Development Cooperation Agency (SIDA)**

*S. Olander*<sup>1</sup>

<sup>1</sup> Swedish International Development Cooperation Agency Sida, Department for Operational Support, Stockholm, Sweden

As the Head of Sida's Evaluation Unit Sven will provide the perspective of Sida, both as a commissioner, manager and user of evaluations. He will speak for 10 minutes and then pass on to the next speaker.

**RT 19 The System for Evaluating International Development Cooperation in Sweden: Is the System Fit for Purpose?**

## **RC 69 - The Perspective of the Expert Group for Aid Studies (EBA)**

J. Petterson<sup>1</sup>

<sup>1</sup> *The Expert Group for Aid Studies EBA, The Expert Group for Aid Studies EBA, Stockholm, Sweden*

As the Managing Director of the Expert Group for Aid Studies, Jan will provide a short introduction to the origin of the EBA, its role and mandate within the Swedish Evaluation System, and how this has developed over the past few years. He will speak for 10 minutes and then pass on to the next speaker.

**RT 19 The System for Evaluating International Development Cooperation in Sweden: Is the System Fit for Purpose?**

## **RC 70 - The Perspective of the Swedish Ministry of Foreign Affairs**

L. Peck<sup>1</sup>

<sup>1</sup> *Swedish Ministry for Foreign Affairs, Department for International Development Cooperation, Stockholm, Sweden*

With his extensive experience of working as an evaluator and evaluation commissioner in international development cooperation for over 25 years, at MFA, Sida and SADEV and as an independent consultant, Lennart Peck has gained an in-depth understanding of the Swedish evaluation system. He will provide the perspective of the Ministry of Foreign Affairs, as well as a more general perspective from his broad background. He will speak for 10 minutes and then pass on to the next speaker.

**RT 19 The System for Evaluating International Development Cooperation in Sweden: Is the System Fit for Purpose?**

## **RC 71 - An Outside Perspective from DEval in Germany**

M. Noltze<sup>1</sup>

<sup>1</sup> *German Institute for Development Evaluation DEval, Competence Centre for Evaluation Methodology, Bonn, Germany*

As a senior evaluator at the Competence Centre for Evaluation Methodology of the German Institute for Development Evaluation (DEval), Dr Martin Noltze will provide a more outside perspective. He will provide an overview of the German evaluation system in international development cooperation, the history, role and mandate of DEval and how the system has changed since DEval started its activities. He will speak for 10 minutes and then pass on to the next speaker.

**RT 19 The System for Evaluating International Development Cooperation in Sweden: Is the System Fit for Purpose?**

## **RC 72 - The Perspective of FCG Sweden, an Evaluation Provider in the System**

J. Lindstrom<sup>1</sup>

<sup>1</sup> *FCG Sweden, M&E Unit, Stockholm, Sweden*

With her extensive experience as a researcher and evaluator, previously in the UK at IDS and now at FCG Sweden providing project and programme evaluations for Sida, DFID and other donors, Johanna will provide the perspective of an evaluation provider in the Swedish system and how this compares to other systems, including the UK. She will speak for 10 minutes and then pass on to the next speaker.

Friday, 5 October 2018  
08:30–10:00

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**RT 20 How to Demonstrate Value for Money? Methods and Trends**

## How to Demonstrate Value for Money? Methods and Trends

M. Branco<sup>1</sup>, A. Richards<sup>2</sup>

<sup>1</sup> [www.mariana-branco.com](http://www.mariana-branco.com), EES Board- EES Emerging Evaluators- EES Social Media, Porto, Portugal

<sup>2</sup> Social Value UK, Research and International Training, Liverpool, United Kingdom

**Rationale:** In this round table different economic analysts will gather to demystify the concept of Value for Money (VfM), present the differences between the current VfM methodologies and discuss VfM trends and challenges. Ultimately, the session intends to provide practical insights that help evaluators understand the meaning, potential and applicability of VfM approaches. To achieve such aim, the round table will create a semi-structured dialog between VfM experts combined with 3 minutes presentations on key methodological implications of VfM, followed by focused and interactive discussions that will promote an open environment for learning and exchange.

### Clarify the conceptual approach (20 min)

This round table will start with a presentation of 3 different definitions of VfM by the moderator. The round table members will then be asked to comment and discuss these definitions, stating the pros and cons of each of them. After a quick debate among the round table members, the audience will be invited to share questions regarding the topic that they would like to see answered by the panel of experts.

### • **Unpack and Reassemble the Value for Money dimensions (30 min)**

The second part of the round table will provide an overview on fundamental VfM dimensions. Round table panelists will be invited to answer the following questions:

- **Unpack:** What is value? Who values? Value for whom? What aspects do we measure? Which part of the value chain are we valuing? How do we capture value (ratio, rating, contribution, narratives)?
- How to combine "value" and "for money"? What are the differences between Economic Analysis, Cost-Efficiency, Cost-Effectiveness: Cost-Benefit perspective (CBA a SROI) and other emerging VfM perspectives?

### Discuss common challenges and applications of VfM approaches (40 min)

The third part of the round table will narrow down the discussion. After creating a common conceptual ground, the panelists will debate key methodological implications of VfM approaches. Each panelist of the round table will have 3 minutes to present what, in their view, is a key challenge, application and recommendation regarding the use of VfM approaches. Following the quick presentations, the room will be invited to a discussion, which will cover the topics below:

- What does the valuation of intangibles benefits and costs entails?
- How to engage stakeholders engagement in the monetization of non-market outcomes?
- What are the differences in the use of Value for Money across sectors (private sector, public sector, NGOs)?
- How can VfM serve as a tool for accountability and adaptive management?
- How to framework questions and criteria for assessing value for money
- What are the implications of VfM in terms of boundaries and causality?
- In which cases is it appropriate to use

**Justification:** Showing the value of organizations in real terms is increasingly important in the social sector. As aid budgets come under increasing scrutiny, aid agencies and development practitioners face a growing need to demonstrate value for money (VfM). This is appropriate and should be welcomed. However, up until now this need had not been matched with appropriate methods to support meaningful VfM assessment, with the prevailing tendency to rely on readily quantifiable measures, even when those don't capture the most important aspects of the change that is being pursued. The term VfM has re-emerged in the vocabulary of public and private spending, particularly, as a tool to cope with the financial crisis, reinforcing the demand for accountability and rigorous impact evaluation. But what do we really mean by VfM? Some argue that VfM is an elastic concept that currently lacks conceptual properties and boundaries, others ask if VfM "isn't just the latest flavour of the month". What is VfM in reality?

#### RT 20 How to Demonstrate Value for Money? Methods and Trends

### RC 73 - The Importance of Engaging Stakeholders in Social Outcomes Monetization

M. Branco<sup>1</sup>, A. Richards<sup>2</sup>, B.C. Carpinter<sup>3</sup>

<sup>1</sup> [www.mariana-branco.com](http://www.mariana-branco.com), EES Board- EES Emerging Evaluators- EES Social Media, Porto, Portugal

<sup>2</sup> Social Value UK, Research and International Training, Liverpool, United Kingdom

<sup>3</sup> Social Value UK and Social Value International, Operations Manager- Assurance & Accreditation services, Liverpool, United Kingdom

Compare social return on investment (SROI) against the different 3 criteria: relevance, completeness and accuracy; Provide examples of different levels of rigour in SROI and how they contributed (or not) to decision making; Discuss how to framework questions and criteria for assessing value for money;

Discuss how outcomes monetization can contribute to promote social resilience and be useful in times of financial, environmental, political and social crises.

#### RT 20 How to Demonstrate Value for Money? Methods and Trends

### RC 74 - Economic Evaluation

E. Sarmiento<sup>1</sup>

<sup>1</sup> Nova School of Business and Economics, Novafrica, Lisbon, Portugal

Distinguish between VfM and Economic Analysis; Discuss how VfM frameworks contribute to more resilient societies and what should be their role in times of economic, political, humanitarian and sustainability crises.

#### RT 20 How to Demonstrate Value for Money? Methods and Trends

### RC 75 - Value for Money Approaches and The Private Sector

K. Siblini<sup>1</sup>

<sup>1</sup> The World Bank, Trade and Competitiveness, Washington DC, USA

Debate technical challenges of the application of VfM in the context of the World Bank Group (e.g.: accountability versus adaptive management; incorporating risk analysis in VfM); Discuss how Cost-Benefit Analysis and Value for Money frameworks can contribute to more resilient societies and what should be their role in times of economic, political, humanitarian and sustainability crises.

**RT 20 How to Demonstrate Value for Money? Methods and Trends****RC 76 - Impact Measurement in the Private sector**

J. Gargani<sup>1</sup>

<sup>1</sup> *Gargani + Company- Inc., President, Berkeley- California, USA*

- Discuss how we should frame questions about impact and evaluative criteria when assessing value for money;
- Debate technical challenges of the application of VfM and monetary valuation (e.g.: usefulness and certainty of estimates in monetary units);
- Suggest new approaches to VfM and contrast them with current trends;
- Discuss the conditions under which VfM approaches contribute to more resilient societies, and their role in times of economic, political, humanitarian, and sustainability crises.

**RT 20 How to Demonstrate Value for Money? Methods and Trends****RC 78 - Value for Money in the Field of Democracy and Rule of Law**

B. Baruch<sup>1</sup>, J. Taylor<sup>1</sup>, E. Dujso<sup>1</sup>, M. Barberi<sup>1</sup>, J. Lonsdale<sup>1</sup>, T. Ling<sup>1</sup>

<sup>1</sup> *RAND, RAND Europe, London, United Kingdom*

- Discuss how to framework questions and criteria for assessing value for money;
- Debate technical challenges of the application of VfM at a macroscale (e.g.: are VfM approaches good for evaluating complex system change);
- Discuss how VfM frameworks contribute to more resilient societies and what should be their role in times of economic, political, humanitarian and sustainability crises.

**RT 20 How to Demonstrate Value for Money? Methods and Trends****RC 79 - Value for Money in the Context of INGO's**

M. Branco<sup>1</sup>, F. D'Emidio<sup>2</sup>

<sup>1</sup> *www.mariana-branco.com, EES Board- EES Emerging Evaluators- EES Social Media, Porto, Portugal*

<sup>2</sup> *Independent Consultant- former ActionAid, Value for Money, Greece, Italy*

- Understand how can VfM move from being an "additional reporting requirement INGOs have to comply with" to a tool for learning and adaptive management;
- Debate technical challenges of the application of VfM (e.g.: whose value counts when we assess VfM), solutions and new trends.

**RT 20 How to Demonstrate Value for Money? Methods and Trends****RC 80 - The Value for Money agenda: A missed opportunity**

F. D'emidio<sup>1</sup>

<sup>1</sup> *Independent Consultant, Independent Consultant, London, United Kingdom*

Value for Money (VfM) was first introduced in the international development sector in 2010 by the then Secretary for International Development, Andrew Mitchell. The concept was imported from the private sector where an investment is considered value for money for a private company if the returns outweigh the cost.

In the not-for-profit sector, while assessing the costs may be relatively straightforward, the value side of the VfM concept is harder to pin down in relation to social change where financial returns are not the driver of the work. It is often difficult to allocate financial value when the work focus is on social, political and human change processes, which frequently cannot be accurately described by quantitative metrics because they involve multiple interrelated issues, often falling outside the market economy.

The VfM agenda was promoted as an invitation, later translated into a requirement, for NGOs to show the public (and the Treasury) what was happening with aid money. Despite a theoretical openness of the donor community to enabling the sector to work out an appropriate understanding of VfM, since 2010 INGOs have faced increasing requirements to demonstrate VfM using the criteria utilized by the private sector, known as the 3Es, Economy, Efficiency and Effectiveness. However, donors as well as INGOs have struggled to demonstrate their VfM to the stakeholders they are accountable to using the 3Es framework and the learning generated across the sector as a result of this has been limited.

For some INGOs, a key question was how they could use the VfM agenda to improve their programmes taking into account the complexity of aid work. In other words, is VfM a new buzz word that relies on strict formulas to demonstrate upward accountability? Or is it a concept that can actually bring about new ways of thinking and working into the sector and be used for downwards accountability as well?

Drawing on ActionAid's experience, this paper illustrates how the VfM agenda may offer an opportunity to move beyond a cost-focussed approach, often translated in a reporting requirement that INGOs must comply with, towards a value-centred tool for adaptive management and programmatic and organizational learning.

To do so, the paper reflects on ActionAid's learning from delivering a VfM Pilot Project, funded by the UK Department for International Development (DFID) through the Programme Partnership Agreement (PPA), aimed to carve its own understanding of VfM and develop appropriate methodologies to measure VfM. The methodology developed engaged directly people living in poverty, predominantly women, and frontline staff and partners, enabling them to assess whether the organization was generating value for the people that it aims to serve.

Friday, 5 October 2018  
11:15–12:45

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

## RT 21 Contribution of Meta-Studies for Building Resilient Societies

### Contribution of Meta-Studies for Building Resilient Societies

*S. Silvestrini*<sup>1</sup>

<sup>1</sup> CEval GmbH, CEO, Saarbruecken, Germany

Starting 1981 with the publication of the JCSEE Program Evaluation Standards, the assessment of the quality of evaluations (meta-evaluations) and the synthesis of their findings (systematic reviews) has become an integrant in evaluation business and research. (For the sake of simplicity in the following the term “meta-studies” is used for both.) Ever since numerous checklists, guidelines and even book publications deal with the methodological foundations of meta-studies and provide practical guidance on how to implement them. Yet, until a few years ago the topic relegated to a niche existence with only few studies being commissioned and even less being published.

Recently, however, it appears that meta-studies come into vogue with a rising demand particularly from the public sector. During the last five years alone, we at CEval conducted a dozen studies, which focused either on (re-)assessing evaluations or their findings, or both. – In the decade before it were none.

While many appreciate this trend as an opportunity to extend their services, it also gives food for thinking about why interest has increased so sudden lately. In any case the reasons for conducting meta-studies remain the same since evaluation has emerged as a means for evidence-based decision-making: scrutinizing the quality of evaluations and the validity, reliability and objectivity of their findings, aggregating these findings, providing accountability, collecting lessons learnt, identifying the benefit of those exercises etc.

The key-question that should be discussed in the round-table session is accordingly: Why is there a rising need for meta-studies? What is their benefit? How are they used and how do they actually contribute to policy-making, and eventually to building resilient societies?

Dr. Silvestrini, will provide a introductory presentation, starting with a brief history of meta-studies in the field of evaluation and the methodological approaches to it, followed by a summary of the key-findings from studies conducted at CEval, which are based on a total of more than 250 evaluations. Finally, some overarching conclusions will lead over to a discussion with four experts, representing the main spheres relevant to meta-evaluation: (meta-)evaluation commissioners (Ilona Mattila, Evaluation Officer, Development Evaluation Unit at Ministry for Foreign Affairs of Finland), policy makers (Michaela Zintl, Director of Evaluation and Portfolio Research at the German Federal Ministry for Economic Cooperation and Development), evaluation practitioners (Lutz Neumann, LBN Strategies, Cologne) and academia (Prof. Dr. Reinhard Stockmann, Director of the Center for Evaluation at the Saarland University, co-founder of the German Speaking Evaluation Society, DeGEval).

After the presentation, each expert should provide her/his perspective on the issues of demand for and use of meta-studies in her/his working field in a brief statement of about 5–8 minutes. In the following Prof. Stockmann will moderate the discussion with questions focusing on the actual benefit of meta-studies for program planning and implementation, evaluation practice, policy-making, and for societal development and resilience at large.

After two rounds of exchange the floor will be opened for a final Q&A with the audience. Eventually, the introductory speaker will close with some concluding remarks, based on the discussion.

**RT 21 Contribution of Meta-Studies for Building Resilient Societies**

**RC 82 - Round-table Contribution**

*L. Mattila*<sup>1</sup>

<sup>1</sup> Ministry for Foreign Affairs of Finland, Development Evaluation Unit EVA-11, Helsinki, Finland

**RT 21 Contribution of Meta-Studies for Building Resilient Societies**

**RC 83 - Round-table Contribution**

*M. Zintl*<sup>1</sup>

<sup>1</sup> German Federal Ministry for Economic Cooperation and Development BMZ, Evaluation and Development Research, Bonn, Germany

**RT 21 Contribution of Meta-Studies for Building Resilient Societies**

**RC 84 - Round-table Contribution**

*L. Neumann*<sup>1</sup>

<sup>1</sup> LBN Strategies, Cologne, Germany

# Posters

Thursday, 4 October 2018  
16:30–17:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

PS S1 Poster Session - Strand 1

## P 01 - Capacity to Do and to Use Evaluation in Community-Based Organizations in Québec (Canada): Development and Validation of a Measurement Instrument

*D. Buetti<sup>1</sup>, I. Bourgeois<sup>2</sup>*

<sup>1</sup> University of Ottawa, Interdisciplinary School of Health Sciences, Ottawa, Canada

<sup>2</sup> University of Quebec, École nationale d'administration publique, Gatineau, Canada

Organizational evaluation capacity refers to the ability of an organization to both do and use evaluations for evidence-based decision-making. Increasing evaluation capacity in an organization can improve its mission-related outcomes since high-quality evidence is systematically informing decisions regarding programs and services. Community-based organizations (CBOs) play a crucial role in promoting health and well-being often where other public services have failed. However, several studies have found that CBOs tend to rely more on anecdotal evidence rather than on the use of evaluation evidence to improve their activities and services. Other studies have shown that CBOs are facing multiple challenges in developing their evaluation capacity including lack of knowledge and evaluation skills; limited access to evaluation expertise, and lack of human and financial resources to conduct evaluations.

In Canada, Québec has more than 8000 CBOs providing various services and is the only province where French speakers are in the majority. Recent reports exposed that CBOs in Québec may experience more considerable challenges in building evaluation capacity in comparison to similar organizations in the rest of the country. In fact, while the province holds the highest number of CBOs in Canada, they are also the most underfinanced and understaffed throughout the country. Measuring evaluation capacity of CBOs in Québec is critical to design a cost-effective evaluation strategy tailored to their organizational strengths and needs. Recently, Canadian research has placed more emphasis on developing measurement instruments to evaluate capacity in CBOs, but no assessment tool has been designed specifically for CBOs in Québec in French.

**Objectives:** This communication aims to present the first empirically tested self-reported instrument in French to measure both capacity to do and use evaluation in CBOs in Québec. The instrument was adapted from an existing tool, the Organizational Evaluation Capacity Self-Assessment Instrument (OECISA), which enables organizations to assess and interpret their evaluation capacity. Methods used to ensure contextual relevance of the instrument for CBOs in Québec were: a systematic review of facilitators and challenges in building evaluation capacity of CBOs in Québec, as well as a two-stage pretest using a panel of experts in evaluation and a sample of community-based workers in Québec.

**Narrative and justification:** Despite being a prosperous country, many groups across Canada are still experiencing disparities in health. CBOs in Québec are vital in the fight against inequalities and in pressuring the government into taking social action. This qualitatively validated instrument is the first to provide CBOs in Québec with a measure of both their capacity to do and to use evaluation in French. This instrument will further assist CBOs and evaluators in (a) planning specific strategies tailored to their current strengths and needs for building evaluation capacity, and (b) assessing their cost-effectiveness in the long term.

Thursday, 4 October 2018  
16:30–17:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

PS S1 Poster Session - Strand 1

## P 02 - Challenges in Evaluating Small-scale Projects with Limited Resources

*S. Fesenko*<sup>1</sup>

<sup>1</sup> OSCE Project Co-ordinator in Ukraine, Project Management, Kyiv, Ukraine

**Rationale:** OSCE Project Co-ordinator in Ukraine has limited resources for implementing a wide number of projects to support development of democratic initiatives in Ukraine. At the same time evaluation of each project is a part of OSCE project management cycle and an obligatory requirement of the donors.

**Objectives sought:** To find new innovative approaches keeping a balance between resources allocated for evaluation of the project and its implementation while getting the most valuable information for decisions and reporting to donors.

**Brief narrative and justification:** Budgets of annual projects implemented with the support of OSCE Project Co-ordinator in Ukraine (PCU) mainly vary between 10,000 EUR and 80,000 EUR. Very often one person manages more than one project with only one assistant assigned for 2–3 managers. Projects are designed to achieve the results in the most efficient way, often overwhelming managers with number of tasks to be implemented. Projects are often focused on the interventions at the level of the whole country, like amendment of the legislation, improvement of national training system, institutional capacity development, facilitation of reforms, etc.

Generally the evaluation practice in the office has two main focuses: to assess effectiveness of events by collecting information at their end; and to evaluate impact at the national level by interviewing key stakeholders at the end of the annual project(s). Such approach is neither creative nor innovative, but does not require investment of additional time and money while providing, to certain extent, information on what has been achieved successfully and what could be improved. Also this practice fits into bureaucracy context where you can provide proof of your evaluation results by uploading primary data to the system, if required.

For activities like series of trainings and workshops conducted in various regions of Ukraine, we recently started to organize annual evaluation conference. The best and most active trainees are invited to share their experience in applying skills and knowledge obtained at the trainings. Such events allow to evaluate relevance of the topics to the audience, effectiveness of methods studied, identify obstacles for use of skills and learn what could be done better. Project managers find evaluation conferences quite innovative and helpful for improvement of future projects.

However, interviews with high level authorities to assess impact of PCU interventions remain to be a point of concern. Since stakeholders are involved in the interventions and realize that partly success of the projects depends on them, they tend to provide too positive/optimistic information.

So, I would like to hear the opinion/ideas of the audience on how to improve our evaluation practice for more creative and innovative, taking into consideration limited human and financial resources.

Thursday, 4 October 2018  
16:30–17:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

**PS S1 Poster Session - Strand 1**

## **P 03 - In Which Way Does Evaluation Help to Make Government More Open and Accountable?**

*V. Kuprij*<sup>1</sup>

<sup>1</sup> National Agency for Civil Service, 1<sup>st</sup> Vice Head of NACS, Kyiv, Ukraine

The proposed presentation will share the methodology of efficiency evaluation of public relations of a ministry, public agency.

Public relations are an important functional component of activities of an executive body. They include, in the first place, receiving feedback, monitoring public reaction to the current policies, ensuring transparency and openness of activities of the state institution. Ukraine's experience demonstrates that today, there are a lot of initiatives and practices of state authorities related to proper public relations. However, viability of these practices very often depends on the leadership of respective bodies. The existing legal mechanisms often identify insufficiency of certain forms and methods of such cooperation. The topic of interaction between governmental institutions and the public shifts rather to provision of information, or more precisely – PR-strategies. The first step toward improving public relations in a governmental institution may be introducing evaluation of its activities in the respective sphere. This will enable us not only to see the real picture but also to identify problems and receive the necessary understanding of what needs to be changed and improved.

By joint effort of representatives of the Ukrainian civil society and the Government, pilot methodology was developed and tested for evaluating efficiency of public relations of an executive body (a ministry, agency, etc.). The object of evaluation is the level of fulfilment of the executive body's function related to interaction with the public in accordance with the Ukrainian legislation and commitments of the Ukrainian Government for implementation of good governance principles.

The evaluation methodology is based on the understanding that public relations take place on four levels and include: (1) provision of information, (2) consultations, (3) active participation of the public in development and implementation of public policy (dialog), (4) partnership. At each of these levels, a governmental institution should ensure that a three-stage process is in place: planning-action-outcomes. At the planning stage, it is evaluated how the executive body develops and approves the necessary action plans, instructions, procedures, and so on. At the action stage – how the executive body ensures adherence to the procedures, which effort is made, which events are organized, respective products created, etc. Then, at the stage of outcomes, it is evaluated what was actually changed, what happened as a result of the effort taken at the previous two stages. The progress at each of the stages and within every stage is measured using a five-point scale from 0 to 4 with the 0.25 increment, where:

0 – 1 – the results are not measured because they are absent, or because only intentions are declared;

1.25 – 2 – insignificant results are available, or individual (one-time) instances;

2.25 – 3 – meaningful results can be seen;

3.25 – 4 – excellent results can be seen that demonstrate sustainability.

Within each of these stages, general assessment is determined that offers an idea about the progress in ensuring efficiency of activities of the executive body in the public relations sphere. Based on the assessment results, recommendations are developed on improving activities in order to achieve higher efficiency.

Thursday, 4 October 2018  
16:30–17:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

PS S1 Poster Session - Strand 1

## P 04 - Strategic Analysis of the Mais Médicos Program: A Performance Assessment

*L.Z. Oliveira<sup>1</sup>, M. Cruz<sup>2</sup>, A. Oliveira<sup>3</sup>, A. Jatoba<sup>4</sup>*

<sup>1</sup> Fundação Oswaldo Cruz - Fiocruz, Presidência, Rio de Janeiro, Brazil

<sup>2</sup> Fundação Oswaldo Cruz, Escola Nacional de Saúde Pública - ENSP, Rio de Janeiro, Brazil

<sup>3</sup> Faculdade CESGRANRIO, Mestrado Profissional, Rio de Janeiro, Brazil

<sup>4</sup> Fundação Oswaldo Cruz, Presidência, Rio de Janeiro, Brazil

The “Mais Médicos” (“More Doctors”, in free translation) is a Brazilian governmental initiative implemented in order to adequate the distribution of physicians, to reform medical education and to increase investments in primary care clinics all over Brazil. The foundations of the programme are in qualifying care, and moving the Brazilian healthcare framework towards a primary care-centered model.

Despite its relevance to the Brazilian universal healthcare system, recent literature demonstrates the Mais Médicos Programme faces several difficulties in accomplishing its strategic goals. Moreover, the medical category resists the implementation of the programme, as it involves Brazilian physicians, and until recently, opinions diverge regarding the capacity of the programme in enhancing the healthcare system.

Thus, this study aims at evaluating the pertinence of the strategic objectives of the Mais Médicos Programme, in order to describe major obstacles in meeting the demands of population, its importance in reorganizing of the system, and its contributions to a reformed primary care-centered framework.

We propose an expert-centered evaluation model composed of mixed methods, based on three kinds of experts as participants, as follows: a) formulators of the Program; b) specialists in human resources and health; c) specialists in primary healthcare, health planning, and management.

We collected data by means of semi-structured interviews and questionnaires with objective questions (using Likert scale responses), in order to assess the programme's problems and their causes. Finally, we employed a Fuzzy Logic model to aggregate the opinions of experts according to previously defined standards (STARFIELD, 2012; MENDES, 2012).

Results indicate that aspects over which the Mais Médicos programme acts are pertinent. Moreover, results also indicate the programme has the means to play an important role to approach the major problems it is supposed to cover. The causes of problems are described adequately, especially regarding the deficit for physicians, and the Mais Médicos programme relates adequately to this aspect. However, results of this study point out that some causes mapped relate to very pertinent structural problems that the Mais Médicos Programme does not approach.

Regarding problems such as difficulties in accessing health services and long waiting times, among the 21 relevant causes, the Programme approaches 57%. Concerning assessing the relevance of the intervention to standards considered necessary to improve quality and structure primary care, 55% interviewed specialists considered non-pertinent, and 22% consider it pertinent.

Even though problems approached by the Mais Médicos Programme are relevant for improving the quality of primary care, it does not take into consideration some causes as its objects of intervention, which reduces the pertinence of the programme. Therefore, we recommend the increase of investments in actions that strengthen teamwork, regulate the training of health professionals, improve the model for hiring physicians, and interact with other Secretariats of the Ministry of Health. Such action might be useful to overcome problems in the practice of primary care.

Thursday, 4 October 2018  
16:30–17:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

PS S1 Poster Session - Strand 1

## P 05 - Advanced Evaluation Tools and Methods for Developmental Interventions: Case Study of Sustainable Impact of a Scholarship Programme on Recipients

M.A. Seabi<sup>1</sup>

<sup>1</sup> *Kagiso Trust, Programmes Management and Researach, Pretoria, South Africa*

In recent times, where developmental agencies seem to be moving one step forward and five steps back due to regression in standards of living despite the social interventions in communities, a need arise for use of a combination or advanced evaluation tools. These are tools required to ensure efficient planning for clear intervention outcomes and impact. In this study an analysis of Kagiso Trust (KT) scholarship programme – Eric Molobi Scholarship Programme (EMSP), indicate that value is created and achieved through rigorous engagement of the scholarship recipients' families, provision of psycho-social and academic support. Bearing in mind that the recipients come from marginalised communities, living far from modern development and resources. These are the poorest of the poor young people, most orphans or living with grand parents and even being child household heads. The study reflected on the application of the social return on investment (SROI) evaluation of the EMPS. A in-depth case study of two recipients of the KT scholarships, already in the labour market, confirms that the rand value of the outcome of the scholarship programme is more than twice the initial financial investment given for the recipients' tertiary education. The engagement with families, and support during academic life proved to yield value beyond the actual monetary value of the intervention/ scholarship. The sustainability of the EMSP development intervention with the respondents of the study using SROI assessment also confirms that the recipients were not sponsoring education of siblings, and providing for their families, as well as a few of those in need of education support in their community. Over the past ten years KT, a non-profit organisation, introduced scholarship programme to support the poorest of the poor high school matriculants or the grade 12 from deep rural areas in contribution to skilled professionals in the areas of science, accountancy and mathematics. This paper confirms that the adjustment of evaluation measures is necessary for emerging challenges encountered with implementation of developmental interventions. The use of SROI to plan and evaluate the implementation of the KT's scholarship programme model (approach), allows for execution of an intervention that has great potential for sustainability and resilience of beneficiaries to realise the trickle-down impact on their families. The social regression requires that any intervention implemented should be forward looking and have intention for sustainable life changing value add beyond the actual financial investment. The application of the SROI forecasting and evaluation tool allows for organisations to deliver developmental interventions with value add beyond the standard use of monitoring and evaluation framework. Ultimately, contributing to sustainable life changing and impactful interventions towards resilient life.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 06 - Evidence on Adolescent Girls' Aspirations in South Sudanese Refugee Camps in Ethiopia

A. Baysa<sup>1</sup>

<sup>1</sup> International Rescue Committee IRC- Ethiopia Program, Community Wellbeing Initiatives, Addis Abeba, Ethiopia

**Rationale:** From 2014–2017, the International Rescue Committee (IRC) implemented the project *Creating Opportunities through Mentoring, Parental Involvement and Safe Spaces (COMPASS)* in three refugee camps (Sherkole, Bambasi, and Tongo) near Ethiopia's border with Sudan. This project was designed to build adolescent girls' life skills and social capital, while improving their access to services. It also engaged the girls' parents/caregivers in monthly sessions to build their knowledge about adolescent development and skills to best support their girls.

**Objective sought:** The IRC partnered with Colombia University (CU) to conduct the COMPASS baseline and endline evaluations, designed to inform evidence-based design and implementation of adolescent girl programming in the humanitarian context. My poster presentation will draw from the external mixed-methods evaluation led by CU, which used a two-arm waitlist cluster randomized controlled trial (RCT) to evaluate the intervention's impact in the lives of participating adolescent girls. The trial was complemented by qualitative data collected from a sampling of girls and parents/caregivers. Additional monitoring data was collected by IRC staff during project implementation, and these findings will be presented alongside CU research results.

**Brief narrative and justification:** At baseline, the responding adolescent girls in the intervention and control groups expressed limited aspirations – nearly half agreed that girls should marry and have their first child before the age of 18. During the curriculum-based life skills sessions, the intervention girls began to express greater imagination, increased confidence, and new ideas about their future through their weekly drawing activities. In a creative workshop held in each camp, the participating girls clearly communicated their visions in statements such as: "A girl wants to go to school", "A girl builds her community", and "A girl can be a doctor". These messages were turned into photographs featuring the girls. On World Refugee Day 2016, the IRC installed an art exhibit in the three camps with these photographs and the girls' drawings. This created a platform for the intervention girls to communicate with their community about the challenges that adolescent girls face, and their visions for a brighter future. It was an important moment for the girls to feel proud of themselves and their accomplishments, to experience being valuable members of their community, and to see their ideas and hopes acknowledged by their parents/caregivers and the community.

At endline, the perception of acceptable age of marriage and first child varied significantly between intervention and control groups. This demonstrates the project's impact on the intervention girls' awareness, self-perceptions, and visions for their future. While there were some changes with the control girls, they were likely influenced through community mobilization, the art exhibit, and/or hearing about session topics from their peers.

The IRC would be honored to present these learnings to researchers, gender activists, funders, policy makers, service providers, practitioners, and wider global audiences, both to highlight the challenges faced by adolescent girls in humanitarian contexts, and to amplify their messages, resilience, and aspirations.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 07 - Potentials and Challenges in the Analysis of Health Governance Policy Networks in Two Regions of the Legal Amazon - Brazil

A. Casanova<sup>1</sup>, L. Giovanella<sup>2</sup>, M. Cruz<sup>3</sup><sup>1</sup> Fiocruz, LASER/DENSP, Rio de Janeiro, Brazil<sup>2</sup> Fiocruz, Daps/ENSP, Rio de Janeiro, Brazil<sup>3</sup> Fiocruz, Densp/ENSP, Rio de Janeiro, Brazil

Governance is a concept widely discussed in the field of public policies but adopted under different hues. Elements of the reforms undergone by the national health policy, such as decentralization in the coordination, commodification in the provision of services and democratization, with institutionalized arenas to expand social participation make us assume that health governance takes place in a policy network, with various players, in relative situation of autonomy and interdependence, that handle their interests depend on the available resources, mediated by rules and by the degree of influence. This study aimed to understand how to shape the policy networks in health regional governance process in Brazil in two different regions of Amazonia, a spatial area that houses important internal differentiation. For this reason, we chose a multiple case study, including a metropolitan region with characteristics of an intensive urbanization process, population density and prevalence of noncommunicable diseases, and another region, with triple international border, characterized by an important indigenous population, free movement of goods and people, significant presence of drug trafficking in the local economy, and prevalence of endemic diseases associated with drug use, accidents and violence. The research sought to combine methodologically a set of techniques such as semi-structured interviews with twenty-seven political actors in the health sector at state, regional and local level, documentary analysis (records of meetings of co-management spaces, planning, and projects) and social network mapping. Existing approaches were combined in the literature for network analysis and political interdependence of standards and institutions of governance policy networks. An analysis matrix was elaborated considering as main dimensions: the actors, the spaces of articulation, formal or informal, the rules and norms that guide the performance of the actors, instruments, strategies and work processes. The mapping of the social network of political actors was done in order to learn other managing networks, not institutionalized within the Brazilian health system, configured to solve the diversity of problems presented to the managers of this region of the country. It had as main categories the frequency of contact, the nature, and type of links and the type of resources mobilized. Research has shown the importance of combining different theoretical frameworks of policy and governance networks considering different research questions in the evaluation of health in regional governance in Brazil. Depending on the search query, a given theoretical shown to be more significant compared to other used. On the other hand, the adopted analysis dimensions (actors, spaces, rules, and processes) led to the understanding of the diversity of policy networks, institutional or not, albeit with important differences in the cases studied, due to historical conditions, social and geographical of the regions. Some networks, known to exist in both regions, could not be adequately revealed, remaining invisible, indicating that this is a limit of the approaches adopted.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 08 - Determinants to Care-Seeking at Public Health Facilities in Rural Bihar and Uttar Pradesh

A. Chandan<sup>1</sup>, M. Baker<sup>2</sup>, A. Mishra<sup>1</sup>, P. Mishra<sup>1</sup>, L. Ray Saraswati<sup>1</sup>, A. Rai<sup>1</sup>, P. Bhandari<sup>1</sup>

<sup>1</sup> RTI International- India, Public Health, New Delhi, India

<sup>2</sup> RTI International, Global Health Division, Washington DC, USA

**Introduction:** India recorded close to a million under-five deaths in 2016 (GBD, The Lancet). National Family Health Survey-4 reported an under-five mortality of 58 in Bihar and 78 in Uttar Pradesh (UP). Currently, most of the rural healthcare seeking takes place through the informal private sector. One of the key strategies for improving child survival is to improve access and quality-of-care provided at public health facilities. We did a survey to understand factors that inform care seeking at public health facilities in rural Bihar and Uttar Pradesh.

**Methodology:** We selected 28 high caseload public health facilities in four districts of Bihar and 36 in six districts of UP. Within each selected facility, we identified children aged 1–59 months with symptoms suggestive of an illness. We did a face-to-face interview with caregivers of these children to understand their background characteristics and pathways to care seeking.

**Findings:** The survey was implemented for around two weeks at each facility. There were two kinds of caregivers: those who sought treatment directly from the public facility without any prior consultation (2,084 in Bihar and 1,179 in UP), and others with a prior consultation with a private provider (451 in Bihar and 514 in UP). Of those with no prior consultation, 56% caregivers in Bihar and 36.5% in UP were from a below poverty line family. Also, 72.3% caregivers in Bihar and 59.4% in UP were from a socially backward caste. Mothers of 51.5% in Bihar and 67.1% in UP were literate. 87.8% caregivers in Bihar and 72.5% in UP visited the public facility for a better treatment. Only, 5.9% caregivers in Bihar and 1.4% in UP chose a public facility because of free medical services. Around 4.4% caregivers in Bihar and 2% in UP went to a public facility on the recommendation of friends or relatives.

Of those caregivers that had a prior consultation, around 65.1% in Bihar and 40.6% in UP visited the public facility for better treatment. 49.4% caregivers in Bihar and 76.4% in UP visited the public facility because of no improvement in the health of their child after treatment from an earlier provider.

**Conclusion:** Majority of the caregivers seeking care from public health facilities are from a poor socio-economic background. Caregivers changed pathways from private to public in search of better treatment. The belief of getting better treatment from the public facility is a much bigger reason to seek care from public facility than the traditional belief of free medical services.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

## PS S2 Poster Session - Strand 2

## P 09 - Impact Evaluation in Global Innovation Hubs

K. Deng<sup>1</sup>, J. Greenwood<sup>1</sup>, A. Parnia<sup>1</sup><sup>1</sup> MaRS Discovery District, MaRS Data Catalyst, Toronto, Canada

**Background:** The growing emphasis on fostering an innovation economy has raised the question of quantifying the impact of organizations such as innovation hubs as a catalyst for growth [AP1]. Often, innovation hubs include multiple components, such as programming for ventures in forms of incubators or accelerators, a physical space for creating connections in the ecosystem, and formal or informal partnerships with public, academic, or private organizations. The impact of each of these components as well as social, environmental, and the economic life of a region has been a topic of investigation in the last two decades. This rapid review [AP2] investigates the methodologies for impact measurement in the academic literature and a sample of impact reports from innovation hubs around the world.

**Methods:** The search methodology included two streams: systematic search of academic databases (SCOPUS, EconLIT, Google Scholar), and purposive sample of select innovation hubs (n=23). Following screening of the articles and reports, 30 studies were included for analysis. The findings were extracted and analyzed to summarize data sources, research design, metrics, and the statistical methods.

**Results:** Our results indicate that the most common research design in the academic studies was the use of cohort comparison of firms associated with a component of an innovation hub (tenants of science parks or participants in incubators) with matched controls. The studies are either panel studies or cross-sectional, and the outcomes measured related to survival, growth, employment, capital, innovation output, and partnerships. Majority of these studies were conducted on several sites (innovation hubs or incubators) Interestingly, within the innovation hubs' impact reports cohort comparison was not referenced for communicating impact. The information reported in these reports included regional economic contribution studies (using input-output modelling), or performance and usage metrics (number of firms and their characteristic and other uses of the innovation hub). The social and environmental impact was not considered in majority of the reports. These reports included metrics on diversity (based on gender and racial or ethnic background) of the visitors, employees, and client firms' employees. In one case metrics related to the environmental impact of the science park was reported.

The findings of this review indicate the potential for innovation hubs and science park to improve on their methodologies and move toward quasi-experimental designs for impact evaluation. Furthermore, academic studies on innovation hubs have been focused on metrics related to economic performance or innovativeness of firms. Future studies have the opportunity to expand these outcomes to investigate the impact of innovation hubs in particular sectors on the relevant desired societal outcomes (e.g. improvement in health outcomes for health technology industry) as result of the establishment of sector specific innovation programming or hubs. [AP4].

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PS S2 Poster Session - Strand 2**

## **P 10 - Impact of Multinational NERICA Rice Dissemination Project on Socio-economic Life of Rice Farmers in Ekiti State, Nigeria**

*M. Dimelu<sup>1</sup>, O. Ojo<sup>1</sup>*

<sup>1</sup> *University of Nigeria Nsukka, Agricultural Extension, Nsukka, Nigeria*

Agricultural intervention programmes are designed to primarily improve the socio-economic and food security status of target groups. The study assessed the impact of Multinational NERICA rice Dissemination Project on socio-economic life of rice farmers in Ekiti State. Multistage sampling technique was used in selecting respondents. Approach to non-experimental evaluations which is the double difference (or "difference-in-difference") (DD) method was employed. It entails comparing a project group with a comparison group both before and after the intervention. The approach assumes that unobserved selection is present and that it is time invariant and the treatment effect is determined by taking the difference in outcomes across treatment and control units before and after the programme intervention. Data were collected from 109 rice farmers through the use of structured interview schedule. Data were analysed using mean statistic, t-test, double –difference estimator and Mann-Whitney U test. The findings revealed that the mean age of the beneficiary farmers was 42 years, while that of non-beneficiary farmers was 41 years. The project significantly impacted on the beneficiaries' access to market and health facilities; possession of refrigerator and wall clock. On the other hand, no significant change was recorded in the size of rice farm, harvested product, income from sale and other areas of their socio-economic life. The results of mann-whitney U test explained that degree of ease of marketing, harvested rice and access to health facilities have significantly different means compared to others. The study recommends active beneficiaries' participation from inception of programmes and supports for extension services to ensure appropriateness of technology and acceptance by target beneficiaries. Also agricultural interventions and technology dissemination should be accompanied with adequate infrastructural and policy supports to promote mass adoption, ensure expected outcome and sustainability of interventions.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 11 - Lost Road: A Dilemma between Independence and Participation of Evaluation in South Korea

*J.D. Jung*<sup>1</sup>

<sup>1</sup> National Information Society Agency, ICT Investment Performance Center, Seoul, Republic of Korea

Currently under the Korea's ICT project evaluation method under which one third of the total projects are selected as evaluation targets every year, each organization performs self-evaluation (benchmarked on Program Assessment Rating Tool), and the Ministry of Strategy and Finance (MoSF) verifies the result. This system pursues outcome-oriented performance management by deducting performance achievement scores if indicators and goals are inappropriate. Moreover, it holds a strong feedback system, under which more than ten percent of the next-year budget is deducted if the final result of evaluation turns out 'poor'.

Each ministry evaluates projects by himself and MoSF (with NIA) checks and coordinates evaluation score of each project. MoSF applies two strategies in this process. One is a continuous conversation and involvement with each organization. The other is an independent and final judgment of ICT projects performance by the check on self-evaluation results. These processes happen a lot of conflicts and noise between MoSF and each organization. In result, nobody is satisfied with ICT projects' evaluation result.

The purpose of this paper is finding some solutions on a dilemma between independence and participation in evaluation process by the case of Korea central government. These solutions are related with the development of objective evaluation model and indicator, an improvement of punitive evaluation (variation of evaluation feedback in addition to budget cut).

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 12 - Effect Decomposition in Prevention Research: Modern Causal Inference Methods

*Y. Kisbu-Sakarya<sup>1</sup>*

<sup>1</sup> Koç University, Independent Evaluation Laboratory- Department of Psychology, Istanbul, Turkey

Preventive interventions are multidimensional, and frequently target many mediating variables to maximize the effects of a program preventing health problems. Mediation analysis is the statistical method used to investigate through which mechanisms a program changes the targeted outcome variables. A mediating variable, also known as the mediator, differs from other third variables such as a confounder or moderator in that the mediator is causally related to the prevention program and outcome such that the prevention program causally influences the mediator, which in turn causally influences the outcome. Mediation analysis is also known as effect decomposition where the total effect of a program is separated into direct and indirect effects of the program on the outcome variable.

Mediation analysis investigates causal mechanisms. Thus, it involves causal inference by definition. However, most current mediation analysis methods rely on assumptions that may not be satisfied for causal conclusions. If the treatment is randomized in a mediation study, causal claims can be made for the effect of the treatment on the mediator. However, randomizing the treatment will not yield accurate causal direct and indirect estimates unless certain assumptions are satisfied since the mediator status is not randomized. Except for the experimental design options such as double randomization, individuals usually cannot be randomized to the level of the mediator because their score on the mediator is a result of their response to the treatment. This poster presentation will review the theoretical foundations of causal inference in mediation, describe methods to estimate causal direct and indirect effects, and report the results of a large study on the performance of the conventional methods and modern causal mediation analysis methods when testing the direct and indirect effects of a program.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 13 - Practical Data Collection Tools for Engaging Stakeholders in Impact Measurement

M. Branco<sup>1</sup>, H.R. Lamas Diogo<sup>2</sup>, T. Culhari<sup>3</sup>, F. Serejo<sup>4</sup>, N. Gomes<sup>3</sup>, L. Ribeiro Queiroz de Araújo<sup>3</sup>

<sup>1</sup> [www.mariana-branco.com](http://www.mariana-branco.com), EES Board- EES Emerging Evaluators- EES Social Media, Porto, Portugal

<sup>2</sup> Ramboll, Social Performance Manager, Salvador da Bahia, Brazil

<sup>3</sup> Voltália Energia do Brazil, Corporate Social Responsibility, Natal, Brazil

**Justification:** Understanding how companies and investors can contribute to the promotion of sustainable development is essential to achieving transformational change. However, recent studies show that although there's a good awareness of the importance of sustainability among companies and investors, there's little consistency in the approach to measuring social outcomes. Whilst assessing a company's financial performance is straightforward, judging its social impact presents more of a challenge.

**Racional and objectives:** This poster will showcase a range of practical data collection tools for interviews and focus groups have emerged from this study. E.g: an emoji physical likert scale that contributes to more effective outcomes measurement in low-income communities; a social impact pizza that when applied in focus groups enables more accurate contribution analysis.

**Project:** SROI analysis of a social project designed by the Brazilian Government, promoted by Voltália (a producer of renewable energy) and financed by the BNDES (Brazilian Bank of Social and Economic Development), who defines a 5% social credit of the total approved investment. In Brazil, from 2008 to 2014, the investment by BNDES in renewable energy (wind farms) achieved R\$10.5 billion of reais. The northeast region, where this project is being implemented, has the biggest potential for wind and solar energy projects and the lowest HDI. The project evaluated consists of a community-managed desalination system that provides (1) access to potable water and (2) a fish-farming effluent treatment station to two northeastern villages of Brazil that were affected by a 5 years drought.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PS S2 Poster Session - Strand 2**

## **P 14 - Promoting Resilience In Partnerships Through The Use of an Appreciative Inquiry Evaluation**

*E. Lawrenz<sup>1</sup>, M. LaRue<sup>2</sup>*

<sup>1</sup> University of Minnesota, Educational Psychology, Minneapolis, USA

<sup>2</sup> University of Minnesota, Office of the Vice President for Research, Minneapolis, USA

This poster will provide practical information for evaluators working in situations involving multiple partners to optimize program outcomes. The objectives are to describe and critique the use of appreciative inquiry (AI) to evaluate a challenging and complex program involving an \$18 million partnership among a state legislature, a public research institution and private industry to provide economic benefit to the state and the world.

The partnership, MnDRIVE (Minnesota's Discovery, Research, and Innovation Economy), was designed to conduct research in four key areas: treatments for brain conditions; advancing industry, conserving our environment; robotics, sensors and advanced manufacturing; and securing the global food supply. AI is a philosophy and approach that engages all participants to induce positive change, identify what is working effectively, and connects that positivity to enact effective change. Through an AI framework, we interviewed four different sets of people (faculty, leaders, students, and industry liaisons/industry partners;  $n = 19$ ) within each of the four MnDRIVE areas.

These interviews produced a model depicting the flow between ideas. In the U.S., funding for research is very competitive and usually delivered on an individual basis. Therefore, this direct, non-specific support was unusual. Although funding was considered critical, it was viewed as a unique opportunity, resulting in more than the same sort of grant-supported research. Participants felt it allowed leadership to hire diverse, new talent and allowed existing researchers to think freely about ideas and take risks. These fresh opportunities and outlooks resulted in inspiration and innovative movement forward, causing people to reach out in new ways resulting in many new connections with other academics, industry leaders and Minnesota's constituents. Supported by these new connections, innovative research flourished. Participants described the research as pushing boundaries, doing "big things" including translational research that likely will impact the economy and benefit the lives of Minnesotans that would not have been possible without the security and long time frame of the funding. Perhaps not surprisingly, this freedom to explore led to a real sense of gratitude on the part of MnDRIVE participants, a sense of pride, and a sense of duty to give back to Minnesota's stake holders.

Through conducting this evaluation, we learned about the unique insights that can be gained through AI. Although the overall evaluation of MnDRIVE also included collection of extensive quantitative data on project outcomes such as people hired, inventions produced, external grants acquired, interactions with industry and outreach activities and these data were valuable in terms of outcomes, they provided little information about how the program had affected the participants or how to optimize outcomes. AI provided the information needed to understand personal effects and how the program could be successful. AI evaluation itself provided benefit to the participants; a great example of evaluation process use. Despite these benefits of AI, this approach would not be sufficient in and of itself to evaluate this sort of initiative. The outcome numbers were necessary.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 15 - The Hopscotch of Evaluation, an Integrative Methodological Approach

*J.A. Ligeró<sup>1</sup>, M. Bustelo Ruesta<sup>1</sup>, N. Muñoz Muñoz<sup>1</sup>, L. Barreto Lopes<sup>1</sup>, A. Porta Lledó<sup>1</sup>*

<sup>1</sup> *Universidad Complutense de Madrid, Máster Propio en Evaluación de Programas y Políticas Públicas, Madrid, Spain*

The proposal of this poster is to show a theoretical and methodological framework through a sequence of 12 phases to design and conduct evaluations. It arises from the need to incorporate an integrative, conscious, holistic, not partial vision of the different approaches and debates taking place in evaluation, with the intention to conduct better and more useful evaluations. The aim of this proposal is to collaborate, by a creative reinterpretation of evaluation methodologies and approaches, to ensure high quality standards of the evaluations conducted by evaluators and to innovate in the ways that evaluation is communicated.

The main goal of the Hopscotch of Evaluation sequence is to reflect about and be conscious of how each step of the evaluation process implies certain decision-making. Based on the idea that this continuous (implicit or explicit) decision-making process defines the type of evaluation that professionals in the field are conducting, our goal is to reinforce the idea that different methods lead to different types of evaluations and to insist on the importance to take account of evaluation's linkages between theory, methodology and praxis.

The Hopscotch of Evaluation is a sequence that involves 12 sets containing questions and practical cases related to the process of design and conduct useable evaluations. Each of these sets correspond to what we consider to be a phase of the evaluation process: 1) motivation to evaluate and analysis of the context; 2) definition of the evaluand, 3) choice of evaluative approaches; 4) operacionalization; 5) search for information, sources and design of techniques; 6) methodological design for impact evaluation; 7) field work; 8) data analysis; 9) interpretation of results; 10) making judgments; 11) giving recommendations; 12) results communication and influence of the evaluation. Gender perspective and ethical issues are incorporated as an iterative reflection crossing all phases.

The poster shows a structured outline in 12 steps or key nodes of an evaluation process and proposes a reflection and a debate about which are the best options depending on each assignment and evaluative context. The result is a sensitive, singular and adapted design of evaluation. Along with the post we are to propose an interactive and jolly exercise to the attendants for moments of leisure.

The added value of the poster is a methodological proposal that integrates different approaches. It can be used as formative tool and as an element to raise debates and reflections on specifics of the discipline. It contributes to generate dialogue, pooling, appropriation, community, consistency of the field of evaluation and transparency through the playful aspect of our profession. Experience shows that one of the most important indicators of the prosperous knowledge construction, practice and vitality of a field, such as evaluation, is to reach a point of maturity where it can be shared as something playful.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PS S2 Poster Session - Strand 2**

## **P 16 - “AcolheSUS Project” at the Dirceu Arcoverde Hospital: use of Co-Management Devices to Promote Changes in Institutional Reality**

*A.R. Lira<sup>1</sup>, D.M. Cavalcante<sup>1</sup>, M.B. Botelho<sup>1</sup>, E.M. Dos Santos<sup>2</sup>, G. Cardoso<sup>2</sup>*

<sup>1</sup> Brazilian Health Ministry, National Humanization Policy Coordination, Brasília, Brazil

<sup>2</sup> Oswaldo Cruz Foundation Fiocruz, National School of Public Health from Brazil ENSP, Rio de Janeiro, Brazil

The experience of implementing AcolheSUS Project at the Dirceu Arcoverde Hospital (HEDA), on the city of Parnaíba, in a state of Brazilian Northeast, has revealed that the use of active methodology and the construction of logical models can provide and produce changes in the hospitals' model of care and management. The projects' proposal is to qualify management practices and care seeking to address the needs of the community that uses the hospital. In this context, it has promoted the reorganization of health service delivery, based on a perspective of comprehensive care in an adverse scenario of devaluation of health professionals, shortage of resources and a vertical and corporate management. Locally, the project has advocated the execution of planning phases associated with an interdisciplinary discussion. Co-management devices and methodologies are employed in order to promote the integration and co-responsibility of both managers and workers in a process of knowledge construction and sharing of ideas. This is based on the implication of all, as subjects of change and protagonists of actions and propositions of improvements. The evidences presented are originated by a collaborative implementation evaluation performed by the Evaluation Laboratory (LASER) of the National School of Public Health (ENSP), Oswaldo Cruz Foundation (FIOCRUZ), along with the Coordination of the National Humanization Policy (CGPNH) of the Secretariat of Health Care (SAS), Ministry of Health, Brazil. Preliminary findings have shown an increase in managers and workers participation and involvement, mostly in the routine care practices, beginning by a critical reflection on the work processes instituted, providing collective improvement solutions between workers and managers. This may reflect a search for practice innovation, with a bigger protagonism of the actors involved and a greater empathy between the hospital workers. They are mobilized for action, demonstrating a rupture in their institutional cultural context of usual apathy and hyposufficiency / deficiencies, fighting against the daily inertia and facing the problems identified in the institution.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 17 - Innovative Methods for Useful Results, by the Center of Excellence for Development, Impact and Learning (CEDIL)

*E. Masset*<sup>1</sup>

<sup>1</sup> London School of Hygiene and Tropical Medicine, Center of Excellence for Development- Impact and Learning, London, United Kingdom

The Centre of Excellence for Development Impact and Learning (CEDIL) develops and tests innovative methods for evaluation and evidence synthesis. Established in January 2017 under a 5-year grant from the UK government's Department for International Development (DFID), CEDIL is administered through the London School of Hygiene and Tropical Medicine and hosted at the London International Development Centre (LIDC). CEDIL has three key principles:

1. Establish a high quality, inter-disciplinary academic centre to innovate in the field of impact evaluation
2. Design, commission and implement impact evaluations
3. Promote the uptake and use of evidence from impact evaluations

In the presentation I will illustrate the work produced by CEDIL so far and consisting of 4 pre-inception papers and 14 inception papers that are setting a research agenda to evaluate development projects in neglected thematic areas (such as, for example, humanitarian assistance, fragile states, governance, infrastructure and the environment) and by developing new impact evaluation methods. CEDIL is a consortium of leading academic institutions in the field of impact evaluation and comprises a multidisciplinary intellectual leadership team of international experts. Examples of paper topics presented include: Fully integrated mixed methods approaches to evaluation; Causal chain analysis; Timely and efficient systematic reviews; New measures, big data and statistical learning; Models of stakeholder engagement; Mid-level theory approaches; Evidence and gap maps. The poster presentation will also illustrate the research agenda that CEDIL will fund through a competitive bidding process in key areas where methods and evidence in international development are lacking.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 18 - Finding the Balance of Power Dynamics Between Donors & Ngos, a Perspective of an Emerging Evaluator in the Global South

C. Mendoza<sup>1</sup>

<sup>1</sup> *Laurues Sport For Good Foundation South Africa, Programmes, Cape Town, South Africa*

**Objective:** The purpose of this research is to critically examine the demand for robust evidence in the sport for development sector. The research questions whether this is suitable and accurate for the sector.

**Design/Methodology/Approach:** The research utilises current sport-for-development organizations in South Africa to analyse the assumptions that adopting strict requirements on monitoring and evaluation will improve practice in sport-for-development organizations. The authors own experiences, working as an emerging evaluator for a donor, are also drawn upon to critique the value of current 'evidence'.

**Findings:** The research illustrates how adopting stringent Monitoring and Evaluation requirements are unlikely to fully realise the intended consequence of improving future practice in the sport-for-development sector. Utilising personal reflections, the impact that the donor power imbalances have on data reported is highlighted, suggesting that this will rarely lead to data that provides a detailed understanding of work in practice. There was a mismatch between what NGOs and donors tell each other and what is potentially possible.

**Originality/Value:** The research utilises previous literature but also provides a rarely available persona perspective on the issue of donor power dynamics in monitoring and evaluation within sport-for-development.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PS S2 Poster Session - Strand 2**

## **P 19 - The Perils of Selecting and Implementing Evidence-Based Programs in Behavioral Health Sciences**

*A. Motta-Moss<sup>1</sup>, M. Maranda<sup>2</sup>*

<sup>1</sup> City University of New York, School of Medicine, New York, USA

<sup>2</sup> Western Michigan University, The Evaluation Center, Kalamazoo- Michigan, USA

The evidence-based movement gained force in the United Kingdom and in the United States in the 1990s. It promised more effective treatment and prevention services in behavioral health. However, selecting and implementing an evidence-based approach may not be as straightforward as promised. There are many considerations and potential pitfalls. Evaluators of behavioral health programs such as substance abuse and HIV prevention are usually the experts in research and evaluation, and they even may have a role in helping program staff and administrators to select the evidence-based model they will use. During formative processes, the evaluators provide program administrators and staff information about how the program implementation is progressing. This presentation discusses the issues in selecting and implementing an evidence-based behavioral health approach. Some of the concerns are: is the evidence compelling? Has it been replicated? Is the program's target population comparable to the original one? How complicated is the approach and what level of staff training would it require for appropriate replication? What resources are available to aid the implementation? How expensive and lengthy would be the necessary training? Examples from two evaluations of behavioral health programs are used to illustrate the benefits and challenges of implementing evidence-based programs. Included is a discussion of evidence-based registries.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 20 - A Review of Scandinavian Donors Evaluation Reports in African Interventions: Methodological Lessons

*N. Ngwabj<sup>1</sup>*

<sup>1</sup> Stellenbosch University, CREST, Cape Town, South Africa

**Rationale:** The outset of program evaluation as a distinct scientific endeavour dates back to the eighteenth century (Rossi & Freeman, 2004). Since then, the evaluation field has grown with the main output of evaluation reports. Based on a literature review and document analysis conducted, the only form of evaluation report research done on the three donors' evaluations are the OECD/DAC peer reviews. The peer reviews focus attention on donor development assistance; expenditure, policy, strategy and challenges. The rationale of this study was that no previous research has specifically focused on exploring the methodological features of evaluation reports for Scandinavian donors. Therefore, no scientific data exists on the methodological features of evaluation reports of DANIDA, NORAD and SIDA in African Interventions. The three Scandinavian countries- Denmark (DANIDA), Sweden (SIDA) and Norway (NORAD) provide extensive developmental assistance and are commended for their humanitarian aid generosity, with humanitarian aid values that exceed the United Nations' 0,7% gross national income. From 2005, the three donors have produced evaluation reports but no scientific data exists on the methodological features of evaluation reports.

**Aim:** to systematically analyse the methodological features of Scandinavian donors' evaluation reports in African interventions to enable DANIDA, NORAD and SIDA, their donor recipients, other development assistance donors and evaluators to understand the nature of evaluation reports being conducted and to identify areas for improvement.

**Objectives:**

- To establish ; who is commissioning the evaluations, in which fields and the kinds of evaluation approaches and methods being used to conduct the reports
- To understand who is involved in the evaluations of Scandinavian donors,
- To determine whom the clients of the evaluations are

**Methods:** The research will follow a systematic review approach on all DANIDA, SIDA and NORAD African interventions published evaluation reports from 2005–2017. The reports will be purposively sampled to answer the stated objectives.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PS S2 Poster Session - Strand 2**

## **P 21 - Qualitative Approaches for Evaluating a Third-Party Payment Mechanism for Obstetric Care in the Culturally Sensitive Context of Northern Cameroon**

Z. Njoumemi<sup>1</sup>, R. Manouore<sup>1</sup>, T. Inrombe Yano<sup>1</sup>, F. Moulioum N.<sup>1</sup>

<sup>1</sup> Health Economics & Policy Research and Evaluation for Development Results Group HEREG- Yaoundé-Cameroon, Monitoring and Evaluation and Operational Research, Yaoundé, Cameroon

The northern regions of Cameroon were affected by the turbulent causes by insecurity, poverty and climate changes that leads to massive internal and external displacement of populations. In these contexts, the voucher scheme project was implemented as a third-party payment system that covers medical and non-medical benefits (emergency transport, accompaniment by a community referent matron, telephones for referrals between clinics) for the pregnant woman from the beginning of the pregnancy to 42 days after her delivery. The voucher is sold to the pregnant woman at US\$12, which represents about 10% of the estimated average total cost of both pregnancy, delivery and post-partum care she would need. This paper presents the gender-based qualitative approaches used for the mid-term evaluation as well as the results obtained and the lessons learned from this experience.

The team of consultants conducted a mid-term evaluation of the project "Launch and Management of the Cameroon Voucher Scheme System" using the following evaluation criteria: relevance, perception, effectiveness, efficiency, deployment capacities, and sustainability / durability. The evaluation method used was essentially qualitative for the collection of quantitative and qualitative data. The team of consultants carried out combined techniques for data collection: desk review, in-depth interviews, group discussions, in-depth analysis of management tools, and stakeholder analysis and briefings. A sample of 78 key informants and 132 FGD's participants was assembled. Content analysis techniques were used for the exploitation and analysis of qualitative data. Microsoft Excel spreadsheets were used for cross-tabulation and drawing charts/graphs. Descriptive statistics were performed for relative frequencies, averages/means, variances and standard deviations.

The mid-term evaluation found that after 30 months of implementation, the voucher scheme for obstetric care covers 11 districts and 81 clinics. More than 81,000 vouchers were sold for a membership rate of 65%. Attendance at clinics has significantly increased. Births with vouchers (55,419) represent 76% of the total. The post-natal consultations for mothers owners of vouchers (37,901) and the post-natal consultations for new-born with vouchers (39,707) represent respectively 86% and 91% of the total. Among the lessons learned, it can be noted that the price setting of the voucher and the implication of the nursing staff in the promotion of the voucher are decisive. The channels of payment of the benefits to the providers (clinics) are to be included in the design of the third – party payment system. The mobilization of community referent matrons increases medically assisted deliveries. Citizen listening through free phone calls contributes to a better case management for women, mothers and their children who are owners and beneficiaries of vouchers scheme in this cultural sensitive context.

The voucher scheme significantly reduced the financial barriers and contributed to the improvement of the quality of care in the accredited clinics. However, the fact that the vouchers are subsidized by the State at 90% raises the problem of sustainability and its articulation with other mechanisms such as performance-based financing and universal health coverage underway in Cameroon.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

**PS S2 Poster Session - Strand 2**

**P 22 - Evaluating Resilience Using Econometric Methods**

*S. Nyamhuno*<sup>1</sup>

<sup>1</sup> USAID Southern Africa, Evaluation and learning, Johannesburg, South Africa

In the development world there are three definitions for resilience

- i. The ability to cope with a shock
- ii. The ability to adjust to environmental changes
- iii. Transformative capacity of an enabling institution
- a. Resilience Index Measurement and Analysis (RIMA)

RIMA was coined by the Food and Agriculture Organization (FAO) as they measured how various communities adjusted to droughts and famine. The index allows FAO to respond appropriately in drought stricken areas.

Where R is resilience

S -stability

SSN-social safety nets

ABS-access to basic services

A-assets

IFA-income and food access

AC-adaptive capacity

It term of measurement, it is important to note that those indicators are not observable but rather latent.

Methodology

- 1. thereby creating a composite index, R.
- 2. To measure the latent variables, one will have to use latent variable literature such as principal component analysis, factor analysis, structural models and multiple indicator multiple causes. Most recent literature uses an explanatory factor analysis, which assumes that residual errors are not correlated with one another and not correlated with the common latent variables. However in food security cannot this cannot be accepted because of high probability of intradimension correlation. Hence structural equations are preferable because they include correlation between residual errors. Whereas these methods need a many computations, they allow for a model calibration until they get a goodness of a fit
- b. Structural Vector Auto Regression (SVAR)

The equation (L)measures general structural shocks

Where  $A(L)$  is a matrix of a polynomial in a lag operator L, as an  $n \times 1$  data vector with and is an  $n \times n$  structural disturbance. Additional the structural vector is uncorrelated serially and  $V$ , where is a diagonal matrix so that the structural disturbances are mutually uncorrelated.

It must be noted that structural parameters and residuals cannot be estimated, we can therefore estimate a reduced form model which is;

Where  $B(L)$  is a matrix in lag operator (L) and  $\text{Var}(\Omega)$ .

. Let denote  $A(0)$  the contemporaneous matrix of the structural form, and the coefficient matrix in  $A(L)$  without, i.e.  $A(L) = A(0) + A_1(L)$ .

To evaluate the difference in resilience between two communities measured by impulse-response functions for a given structural shock.

, let us consider the structural moving average representation for the vector of variables of interest :

$$= C(L)$$

The impulse response of the structural shock  $j$  on the variable  $i$  for a given country can be represented as:

$$= (L)$$

Where; where

Differences in the individual impulse are computed for a structural shock  $j$  for country  $l$  and  $m$ , namely; –and confidence intervals are calculated by bootstrap procedure as presented earlier.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 23 - Conceptualising Behaviour Science Within Sustainable Perspective

*K. Odhiambo<sup>1</sup>, M. Dianah<sup>1</sup>*

<sup>1</sup> *University of Nairobi, Psychology, Nairobi, Kenya*

Rational Human behaviour and relation to ecology is important more so the way people behave and respond more so as a result of the effects that arise from mitigations such as Climate Change. In SDGs this result in mitigation and adaptive survival traits. This could be negative or positive depending on the extent to which the behaviour response leads to depletion of the environment. Objective: Is to introduce evaluators to the behaviour perspectives and their link to sustainable circumstances Rational and Justification: It is important to understand the link between behavior and environmental factors and effects. The issue is there is not much knowledge in this area as it is new and the world is only now turning to behaviour science to enhance development response more so sustainability. This section will attempt to conceptualise and introduce the audience to this link. There is need to determine review the arising psychological adaptive measures such as appraisal of the situation; cognitive analysis of thought processes; disengagement measures, defensive responses, and emotional regulation. There is also issues such as proximity, exposure, sense making, social construction, and risk as well as collective efficacy This is based on the fact that behaviour science perspective is indispensable for increasing understanding of the drivers of SDGs. Behaviour perspectives that arise consist of intra-individual parameters and processes. First behaviour will be defined, behaviour being a psychological attribute and generally referred to as the "the science of human behaviour and mental processes." Thinking and feeling is included here. It has its roots perception, attitude, values and beliefs. Each attribute results in how we response during on-going environmental experience. This could result in conditional habits or behaviourism; learned behaviour or social; thinking and acting by reason or cognition; and taking responsibility for the actions we decide on or humanistic. The term survival comes to mind here. The behaviour aspects that arise addresses the following parameters: appraising situations, affective response or emotions, cognition or analysis through reasoning, disengaging or separating as well as defensiveness or self justifying. This brings with it collective sense making or shared and social comparison or communality. The behaviour scenario that arise are many regarding peoples environment. The presenter will show this by exhibiting what happens. For example: "...there is at such a time threat appraisal arises \_ what the affected should do and if it will make a difference..." "... There is also Risk Perception \_ which arises from dread, uncertainty catastrophe, prior vicious experience and psychological distance..." The presenter will also show that behaviour is complex and be understood from a theoretical basis and fundamentals of programme logic and form. This paper is part of the PANEL Presentation of The class of 2017 who took a course on "Programme Evaluation and Development" \_ Taught and guided by Dr. Karen Odhiambo\_ University of Nairobi. This was from the classroom Task on SDGs and M&E.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

## PS S2 Poster Session - Strand 2

**P 24 - Patient-Provider Communication Related to Appropriate Drug Administration: Key to Improving Treatment Adherence***A. Rai<sup>1</sup>, M. Baker<sup>2</sup>, A. Mishra<sup>1</sup>, P. Mishra<sup>1</sup>, L.R. Saraswati<sup>1</sup>, A. Chandan<sup>1</sup>, P. Bhandari<sup>1</sup>*<sup>1</sup> RTI International-India, International Development Group, New Delhi, India<sup>2</sup> RTI International, Global Health Division, Washington DC, USA

**Background:** Diarrhea and pneumonia accounted for 32% of the 0.9 million under-five deaths reported in India in 2016. Most of these deaths were related to inappropriate case management, especially, poor drug adherence. We undertook a study to understand treatment compliance and communication on drug administration (essentially the three parameters: dosage, frequency, and duration) of oral rehydration salt (ORS), zinc, and amoxicillin by health providers at selected public health facilities in Bihar and Uttar Pradesh (UP).

**Methods:** We used an observation checklist, adapted from the World Health Organization's facility-based integrated management of neonatal and childhood illnesses tool, to assess the quality of patient-provider communication related to drug administration in 64 public health facilities in Bihar and UP. Patients were children aged 1–59 months that were diagnosed with diarrhea and pneumonia. Additionally, we undertook a face-to-face follow-up interview with caregivers on the 15<sup>th</sup> day of the survey to ascertain their drug adherence and clinical outcome.

**Results:** The study was conducted for around two weeks in each health facility. In both Bihar and UP, 347 children were diagnosed with diarrhea and 52 children were diagnosed with pneumonia. Of those diagnosed with diarrhea, 42.4% were prescribed ORS and 11.8% were prescribed zinc, and of those diagnosed with pneumonia, 23% were prescribed amoxicillin. Of those that were prescribed ORS, 40.8% were informed about the correct dosage, 36.7% about correct frequency, 26.5% about correct duration, and 42.9% about the correct ORS preparation. 21.1% were informed about all the three drug administration parameters. Similarly, of those that were prescribed zinc, 70.7% were informed about the correct dosage, 58.5% about correct frequency, 39% about correct duration, and 41.5% were informed about all the three parameters. Of those that were prescribed amoxicillin, none were explained all the three parameters, whereas 33.3% were informed about the correct dosage, 58.3% about correct frequency, and 8.3% about correct duration. We followed up caregivers on the 15<sup>th</sup> day and self-reported drug adherence was 96.3% for ORS, and 100% for both zinc and amoxicillin.

**Conclusion:** Availability of drugs at public health facilities in the absence of appropriate treatment counselling will not improve the clinical outcome related to childhood diarrhea and pneumonia. Although, we found a high drug adherence, incomplete information on dosage, frequency, and duration is associated with ineffective treatment. We need to strengthen the skills of facility providers on treatment counseling to ensure drug compliance and to achieve the desired effect of the medications.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 25 - Piloting of Using Quick Respond (QR) Code and Imaging Technology to Support Routine Field Monitoring and Evaluation Activities

A. Rujiranupong<sup>1</sup>, T. Sinsomboonthong<sup>1</sup>, K. Maddick<sup>2</sup>, N. De Loughry<sup>3</sup>

<sup>1</sup> Right To Play Thailand Foundation, Monitoring & Evaluation and Learning Unit, Bangkok, Thailand

<sup>2</sup> Right To Play Thailand Foundation, Refugee & Migrant Program, Mae Sot, Thailand

<sup>3</sup> Right To Play Thailand Foundation, Director for Thailand Country Office, Bangkok, Thailand

Right To Play Thailand had conducted piloting of adaption of using the Plickers application on iOS-based mobile device together with a quantitative questionnaire tool to support after event assessment. The Plickers tool is QR code technology and image capturing application, which was initially invented for the purpose of classroom-based administration and assessment. The application was adapted for assessing the quality of our "Friendship Games" event to capture the participants' perspective of the following aspects: the quality of the conducted event, expectation of participants, and satisfaction of participants respectively. The results of piloting of the tool/application are focused on (i) feasibility of using tool as alternative field data collection tool, (ii) satisfaction of using the tool, and (iii) use case in reality and limitation of the tool. It was obviously proven that the Plickers application can easily adapted for using as field data collection tool. There were some limitations found such as (i) difficulty to capture photos via mobile device, especially for when conducting interviews in a large group, (ii) there is a need to practice both use of the tool as well as in questioning and class/event management, and (iii) we were unable to directly export the data to a spreadsheet and word-processing formats/tools.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

## PS S2 Poster Session - Strand 2

**P 26 - Impacting Social Norms Around Early Marriage Through Mass and Community Media: Lessons From Rural India**A. Saha<sup>1</sup><sup>1</sup> Catalyst Management Services Private Limited, Results and Insights, Bangalore, India

Child marriage has been recognised as a major challenge facing developing countries and some estimates suggest that over a third of women aged between 20 and 24 were married before they reached the age of 18. Nearly half of India's brides are married before 18 and despite decline in incidences of child marriage from 47% in 2006 to 27% in 2016 (NFHS-3, NFHS-4), there remains significant inter-state and rural-urban disparities. Given the scale of the problem, Breakthrough, a global human rights organisation intervened in three districts of the states of Jharkhand and Bihar beginning in 2012. The interventions utilised a combination of stakeholder engagements of local leaders, frontline workers etc., community mobilisation and mass media strategies to reduce the incidences of child marriage. The intervention is a Randomised Control Trial that utilised geographical administrative units as the Primary Sampling Units. While the endline evaluation is scheduled for 2019, a midline evaluation was undertaken to understand the progress of the intervention and what are the components that have contributed the most. The ongoing monitoring exercises have also provided crucial insights and helped improve strategies of delivery of the interventions. To monitor outcomes, we used a modification of the Lot Quality Assurance Sampling method given the very limited time and resource available, incorporating elements of social psychology in the surveys. Midline findings suggest that age of marriage has indeed been moving but over a longer period of time and has been observed in both treatment and control groups. As a result, the impact could not be attributed to the intervention. While there are certain infrastructural and resource specific concerns such as lack of schools that is beyond the scope of the intervention, change in social norms regarding marriage and education of girls have resisted change. Data from monitoring show that although community mobilisation activities generate discussions and dialogue in the community, it's fairly limited. The penetration and reach of mass media such as television/radio in rural areas is barely in half of the households but the discussions to impact societal level changes still needs to be pushed further. Stakeholder trainings were a crucial component of the treatment package but their convergence at the community level to own up the problem and address it at their level is still lacking. However, what we clearly see is that knowledge and awareness has diffused in the community but that has not been enough to change norms, though individual behaviour level changes have been captured as anecdotal case stories from the field. Data suggests that there's a case of pluralistic ignorance and we suggest strategies that we believe can work in rural resource poor settings.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 27 - Mixed Methods to Evaluate of the Effectiveness of Task Sharing to Community Health Workers – Evaluating: Preparedness, Safety and Effectiveness

A. Teklu<sup>1,2</sup>, A. Seid<sup>3</sup>, T. Demlew<sup>3</sup>

<sup>1</sup> Ethiopian Evaluation Association, GA, Addis Ababa, Ethiopia

<sup>2</sup> MERQ Consultancy, Addis Ababa, Ethiopia

<sup>3</sup> Ethiopian Public Health Association, Addis Ababa, Ethiopia

**Introduction:** The government of Ethiopia has decided to share the task of provision of Long Acting Reversible Contraceptives to the community level health workers known as the health extension workers (HEW) – community health workers with formal training to provide predominantly preventive services. The program was initiated in 66 health posts as a pilot program. This is believed to address challenges related to access – a major challenge for the 70% Ethiopian women who reside in rural Ethiopia. There was no local evidence showing preparedness, safety and effectiveness of this initiative. This evaluation was done after the program was initiated.

**Objectives:** The objectives of the evaluation were to determine adequacy of the space and the available equipment and supplies at the health post level, to determine the level of knowledge and skills of the providers, to explore the outcomes (success and failure) on the users and determining its safety (reported complications)

**Brief narrative and justification:** Health post level provision of LARC is a new consideration in Ethiopia; thus defining the package properly and starting in limited health posts was needed, before consideration of this initiative for large scale implementation. It was started in 66 health posts in four regions. The program was underway for 8 months. except for some program level and training-related data, there was no user and provider-level baseline data which was taken when the task-sharing pilot program was started.

**Methods:** A post-only study design which employed mixed methods approach was used. Since the health posts were only 66, all were targeted to assess adequacy of space, availability of equipment and supplies, we also assessed ability to function in the absence of power and water supply, infection prevention practices and service availability using assessment checklist which used the WHO building blocks as a reference; and qualitative interviews of the health extension workers.

Knowledge of the health extension workers were assessed using the standard, self-administered questionnaire and their skills were assessed using the objective, structured Clinical examination (OSCE) using anatomical models and the examiners were master trainers of LARC.

Effectiveness of the program was assessed by taking a random sample of women who received LARC by the HEW at the HP-level from the registers at the health post level and by going to the community, finding their household and doing interviews which were done by independent interviewers. The follow up questionnaire focused on satisfaction, discontinuation, any complications (misplacement, infection, and pain), any failure (occurrence of pregnancy), incidence of side effects, knowledge of the users about the method they are using and their attitude towards getting the LARC services at the health post level. Safety was assessed using facility reports and reports from the users including perforation, excessive bleeding and death. **Data collection and Analysis:** manual data collection was used by independent data collectors. Data quality assurance methods were put in place. Parallel analyses of the qualitative and quantitative data were carried out and convergence was done at the results presentation and discussion level.

**Results and Discussion:** the set up and the human resource requirements were found to fulfil the minimum requirement, but there was room for improvement. Discontinuation was at 5.4% and incidence of pregnancy was only 2 out of 568 and there were no major safety concerns identified.

**Conclusion and Recommendations:** a post-only design with no comparison which employs mixed methods approach to evaluate such complex programs provides vital evidence and gives clearer picture on the different aspects of the program. Use of qualitative approaches to describe the implementation and care experience has created clarity. We strongly recommend use such methods to evaluate complex programs with limited baseline data.

Thursday, 4 October 2018  
16:30–17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 28 - The Periodic Table of Evaluation

S. Vaca<sup>1</sup>

<sup>1</sup> *Universidad Complutense De Madrid, Evaluation, Madrid, Spain*

The Periodic Table of Evaluation seeks to present a (pretty exhaustive) catalogue of the different evaluation choices: all (most of) the different alternatives and variation that can be presented in an evaluation, classified by types. It covers Paradigms, Purposes, Objectives, Criteria, Approaches, Designs and Methods. Finally, it also includes some elements that are contextual and inherent to the field of evaluation.

It has two main purposes: 1) broaden the espectrum of choices that commissioners and evaluators have in mind when designing an evaluation; and 2) represent a catalogue of all the things an evaluator can/should/could master, in order to become a better evaluator who make the best possible choices for each particular evaluation context.

Thursday, 4 October 2018  
16:30–17:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PS S3 Poster Session - Strand 3**

**P 29 - From the Tree to the Forest: Reviewing Alkin's Approach**

*E. Álvarez<sup>1</sup>, A. Fowler<sup>1</sup>, O. Mateo<sup>1</sup>, A. Barajas<sup>1</sup>, A. Porta<sup>1</sup>, A. Landa<sup>1</sup>, L. Barreto<sup>1</sup>, A. de Francisco<sup>1</sup>*

*<sup>1</sup> University Complutense of Madrid, Centro Superior de Estudios de Gestión, Madrid, Spain*

Evaluation is an interdisciplinary field of work in which different professional and academic points of view, paradigms and interests converge. Thus, people with all kinds of professional backgrounds are interested in evaluating projects, programs and policies (among other interventions), often with different objectives in mind. This feature of the profession has inevitably meant that, throughout the last few decades, evaluation academics have been discussing a range of important topics, such as, what is evaluation, what techniques should be used for better results, and what are the main uses and purposes of evaluating interventions. As a consequence, several debates have arisen, as evaluation academics have been researching and theorizing, and new schools of thought have appeared.

An internationally renowned academic and expert in evaluation, Marvin C. Alkin, designed in his edited book *Evaluation Roots: A wider perspective of theorists' views and influences* an 'evaluation theory tree', where he groups different scholars that have made relevant contributions to the field, according to which dimensions of evaluation they consider more important, as interpreted by Alkin. In our role as students of a Master's Degree in Evaluation of Public Policies and Programs, and inspired by Alkin's proposal, we have identified several debates between evaluation scholars that are key to the development of the field. Therefore, as students and as a consequence of extensive class discussions on the subject, we were interested in the idea of proposing our own trees (one for each debate). In doing so, we hope to contribute to the discourse on evaluation theory by classifying some of the most important scholars in the tree branches according to our assessment of their positions in each of the debates.

The final product will be a poster, the objectives of which are as follows: Firstly, to create a visual, useful representation of the main debates in evaluation that can be used for educational purposes; Secondly, we would like to bring the debates to the delegates' attention and set up a space in the conference to discuss the different positions of the participants in the evaluation forest; Thirdly, to create a first design of an evaluation forest that can be improved for other students (of our master's degree or others), adding more authors and/or more debates, in order to collectively create a practical and useful visual representation of some of the most important topics being discussed in our field of work.

Thursday, 4 October 2018  
16:30–17:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

PS S3 Poster Session - Strand 3

## P 30 - Guided Technical Visit: Bridge Device Between Multiprofessional Networks and Users, for the Mobilization of Knowledge in an Institutional Learning

D.M. Cavalcante<sup>1</sup>, M.B. Botelho<sup>1</sup>, H.F.D.M. Monteiro<sup>1</sup>, E.M. Dos Santos<sup>2</sup>, G. Cardoso<sup>3</sup>

<sup>1</sup> Ministério da Saúde do Brasil, Coordenação Geral da Política Nacional de Humanização, Brasília, Brazil

<sup>2</sup> Oswaldo Cruz Foundation Fiocruz, National School of Public Health from Brazil ENSP, Rio de Janeiro, Brazil

<sup>3</sup> Fiocruz, Escola Nacional de Saúde Pública do Brasil ENSP, Rio de Janeiro, Brazil

This work aims to present an innovative intervention experience, "AcolheSUS Project", in a maternity hospital, in northeastern region of Brazil, one of the most deprived areas of the country. The hospital is a state reference for expected and high-risk pregnancy, but presents high maternal mortality rates. The "AcolheSUS Project" aims to promote the improvement of health production and promotion practices, as a priority strategy of the Humanization Policy of the Brazilian Unified Health System (SUS). We employed the triple inclusion "method", including different actors, collective analyzers of conflicts and networks. The evidence presented is a result from a collaborative implementation evaluation developed by a technical cooperation between the Evaluation Laboratory (LASER) of the National School of Public Health (ENSP), Oswaldo Cruz Foundation (FIOCRUZ) and the Coordination of the National Humanization Policy (CGPNH) of the Secretariat of Health Care (SAS), Ministry of Health, Brazil. One of the main devices used to undertake a situational diagnosis and help in the construction of the work plan was the guided-visit, which permits us to walk through the itinerary performed by the patient within the health facility. This device seeks to problematize, from health professionals, workers and managers perspective, patients' access to the different units and types of care inside the facility. It also enables a transposition in the pedagogical process, since it mobilizes those involved in order to analyze their own insertion in the process, and the relationship established with other professionals involved in the daily work activities. This permit viewing possible processes of change, surpassing, thus, the mere supervision of activities or monitoring data. The first guided visit was held in 2017, when we identified and problematized among the participants the different critical points in the patient's trajectory inside the facility. These were: care fragmentation activities, absence of organizational flows and lack of joint responsibility in the professionals and general workers practices. In January 2018, a second guided tour was held, which triggered changes in work processes (alliances for the search for solutions). The priority problem selected then was the fragmentation of care activities, mostly in the "Joint Housing" unit (place where mother and baby stay together after delivery). This problem served as a starting point for the organization of a set of workshops for the construction and modeling of an intervention work plan and an evaluation proposal. As lessons learned, periodic guided visits enabled to outline unexplored perspectives, not usually visualized with traditional direct observation or monitoring tools. It also stimulated health workers protagonism and valued their collective creativity in the constructing of strategies to reinvent conventional modes of actions.

Thursday, 4 October 2018  
16:30–17:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

PS S3 Poster Session - Strand 3

## P 32 - Conceptualising Behaviour Science Within Sustainable Perspective

K. Odhiambo<sup>1</sup>, M. Dianah<sup>1</sup>

<sup>1</sup> University of Nairobi, Psychology, Nairobi, Kenya

Rational Human behaviour and relation to ecology is important more so the way people behave and respond as a result of the effects that arise from mitigations such as “Sustainable Development and Climate Change. In SDGs this result in mitigation and adaptive survival traits that a have behavior inbuilt. This could be negative or positive depending on the extent to which the behaviour response leads to depletion of the environment. This shift requires an evaluation system that create the logic required to evaluate. Thus it is important to raise the M&E issues. This requires being able to identify the behavior components and the schemas of measure. There is need to address the arising psychological adaptive measures of behavior

**Objective:** Is to introduce evaluators to the behaviour perspectives and their link to sustainable circumstances Narrative and the way in which M&E plays a role in ensuring a theory of logic.

**Justification:** It is important to understand the link between behavior and environmental factors and effects and how M&E can help link the synergies that arise *and how the behavior schemas can be measured*. The issue is there is not much knowledge in this area of behaviour science more so sustainability and the link to M&E. This section will attempt to conceptualise and introduce the audience to this link. This is based on the fact that behaviour science perspective is indispensable for increasing understanding of the drivers of SDGs. Behaviour perspectives that arise consist of intra-individual parameters and processes. First behaviour will be defined, behaviour being a psychological attribute and generally referred to as the “the science of human behaviour and mental processes.” Thinking and feeling is included here. It has it's roots\_ perception, attitude, values and beliefs. Each attribute results in how we response during on-going environmental experience. This could result in conditional habits or behaviourism; copying, thinking and acting by reason or cognition; and taking responsibility for the actions we decide on or humanistic. The term survival comes to mind here. The behaviour aspects that arise addresses the following parameters: appraising situations,, affective response or emotions, cognition or analysis through reasoning. This brings with it collective sense making or shared and social comparison or communality. The behaviour scenario that arise are many regarding peoples environment. The presenter will show this by exhibiting how these can be measured and challenges that arise in evaluating. The presenter will also show that behaviour is complex and be understood from a theoretical basis and fundamentals of programme logic and form. This paper is part of the PANEL Presentation of The class of 2017 who took a course on “Programme Evaluation and Development”\_Taught and guided by Dr. Karen Odhiambo, University of Nairobi. This was from the classroom Task on SDGs and M&E.

Thursday, 4 October 2018  
16:30–17:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

PS S3 Poster Session - Strand 3

## P 33 - Evaluation of Monitoring and Evaluation Processes Institutionalization Related to National Technical Cooperation of the Ministry of Health with Federal Institutions

J. Kabad<sup>1</sup>, E. Moreira dos Santos<sup>1</sup>, A.C. Figueiró<sup>1</sup>, D. Abreu<sup>1</sup>

<sup>1</sup> Fundação Oswaldo Cruz FIOCRUZ, Laboratório de Avaliação de Situações Endêmicas Regionais LASER, Rio de Janeiro, Brazil

This paper presents an evaluation research on the institutionalization of monitoring and evaluation processes in the International Technical Cooperation Coordination (CPCI), linked to the Department of Health Economics of the Ministry of Health in Brazil (MS), responsible area for management and monitoring of the international and national technical cooperation. As a participatory and iterative evaluation, approach involved stakeholders at all stages and made deliveries of immediate use by the applicants. It aimed at evaluate managing of the national MS technical cooperation with other Federal Institutions, carried out through the Terms of Execution Decentralized (TED). The TEDs are legal instruments that allow the direct transfer of resources between federal institutions in the country with or without intermediary support foundations. This mechanism facilitates to perform complementary functions to the MS in the areas of research, teaching, production of health inputs and innovation. In the current Brazilian context, the TEDs have made possible arrangements for solving public management problems emerging from changes in the state, increasingly directed to a business market function. Two questions guided this assessment: What are the benefits of institutionalizing monitoring and evaluation to strengthen public accountability? What path has been taken towards institutionalizing the evaluation and management of health knowledge in the organization studied?

The evaluative research carried out in the period from 2009 to 2015 involved the study of 1,607 TEDs. The theoretical basis considered the TEDs as a mechanism of technical cooperation, which can be referenced to the knowledge management approaches, especially those anchored in the production and reproduction of socio-technical networks. Immersed in Brazilian public sector, it understands the principles of public administration, especially those of the professionalization of management, accountability and learning about the machinery of government. Research methods involved bibliographic review, documentary analysis, descriptive exploratory analysis of databases available through "BIG DATA" and the study of two cases, institutions that added 42% of the TEDs in the period (case 1 = 477/1607 and case 2 = 201/1607) and which received higher sums of resources during the study period. Initially the exploratory analysis searched to describe the TEDs accord to the purposes and related to the domains of financing; the second stage was a deepening analysis of the two cases. The case studies involved a direct observation of the functioning and process flow in each institution, documentary analysis and semi-structured interviews with key actors involved in the networks to identify significant actors, interests and controversies, interactions, mediations, inscriptions and consequences.

The results to be presented correspond to: 1) Modeling of the CPCI functions; 2) Description of the volume of investments in technical cooperation in the period from 2009 to 2015 and the typology of TEDs purposes; and 3) understanding of the flows and networks of organizations and actors built around cooperation with the MS for the two cases analyzed. In summary, the results point out arrangements and interactions to be considered in this critical moment of cuts in the financing of public social policies in Brazil.

Thursday, 4 October 2018  
16:30–17:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PS S3 Poster Session - Strand 3**

## **P 34 - A Multilayer Mechanism to Maximize Results and Embed Evaluations**

*C. Karasi-Omes<sup>1</sup>*

<sup>1</sup> LuxDev, Evaluations- Quality and Knowledge Management Department, Luxembourg, Luxembourg

Evaluations are often seen from the field as an administrative obligation. From a commissioner's perspective, one would like to have all kind of solutions out of evaluations: solutions for decision making, prospects for the future, results achievement from past intervention; while guaranteeing evaluation independence and transparency. Organizing embedded useful evaluations is often a difficult task for those who are in charge of designing evaluations.

The Luxembourg Development Agency, LuxDev, put in place a multilayer mechanism improving acceptance of evaluation processes, results communication, thus leading to improved embedding of evaluations.

First, a big effort is made on the planning and the evaluation design. To improve the evaluation's usefulness, the link of the evaluation process to important decision making or learning moments during implementation is crucial. Secondly taking into account the perspective of the field teams, including their counterparts, in the drafting of the terms of reference, leads to a higher acceptance and ownership of the evaluation process as a whole by all implicated parties as well as in particular the findings and recommendations. While this way to design is very participatory, the evaluation department keeps the lead in all decisive steps (finalization of the terms of reference, expert recruitment, report final acceptance) in order to guarantee the process' independence and transparency. This level is the operational field level.

From the headquarter perspective, the usefulness comes from a close implication of the people in charge of monitoring programs and the sector or thematic experts. Implicating their needs as evaluation questions and inviting them systematically to the briefing and debriefing sessions, result in the use of the evaluation lessons learned and conclusions in the organizational learning, thus the knowledge capital of the Agency. This level constitutes the organizational learning level and link with knowledge management and capitalization processes.

While technicians in the field and experts at headquarters are very interested to go in detail through the report and/or participate to two-hour debriefing sessions, the management level or representatives from ministries are very busy and have less time to "digest" evaluation reports. Report summaries, even though already more compatible for a director's or donor's reading, do not often highlight the major points well enough and are often "nicely" written as they are published on the websites for a wider public. In order to enlarge the lessons learning to the higher management level, a mini debrief has been introduced at LuxDev taking place just before the debriefing. The objective is to concentrate during 20 minutes on the major issues that are important for the management level or the donor. This management level improves also the evaluations internal visibility.

Through the multilayer mechanism, LuxDev maximizes internally evaluation results dissemination, internal learning as well as the link with the management, thus strategic decision making and internal visibility.

Thursday, 4 October 2018  
16:30–17:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

PS S3 Poster Session - Strand 3

## P 35 - Evaluation of Good Projects: Looking for Opportunities and “Power Results”

C. Karasi-Omes<sup>1</sup>, A. McCarty<sup>2</sup>, A. Jacquemin<sup>3</sup>

<sup>1</sup> LuxDev, Evaluations- Quality and Knowledge Management Department, Luxembourg, Luxembourg

<sup>2</sup> Mekong Economics, Consultant Company, Hanoi, Viet Nam

<sup>3</sup> LuxDev, Project VIE/033, Hue, Viet Nam

We tend to seek lessons from failures rather than successes. Evaluation of good projects is therefore somewhat of a formality, with success attributed casually to many factors. In this paper we take one such successful project and show how more powerful insights into lessons to learn can be developed.

The Luxembourg Development Agency, LuxDev, mandated Mekong Economics for the final evaluation of a project for local development and adaptation to climate change impacts in the lagoon area of TT Hue province, Vietnam. The project was assessed as very successful at midterm and did not face any major challenge since then. The challenge was thus to find opportunities for interesting findings maximizing evaluation results.

The first opportunity was the fact that the baseline was done in more districts than the project thereafter covered. The baseline household (2014) survey interviewed a random sample of 904 households in 13 communes of 29 project target communes, and 196 households in 3 communes outside of target districts. The endline household survey (2017) interviewed the same 904 households as during the baseline, and it was decided to re-interview the 196 households as a control group. This randomized control trial revealed among other variables that the control group was marginally wealthier pre-project (2013), yet much poorer than the treatment group after the project (2017). Although the control group was small, this was a fascinating insight into trying to answer the counterfactual question of attribution – for an agency normally content with before-and-after logframe indicators.

A second opportunity emerged when Mekong Economics identified (and labelled) “Power Results”. These were defined as results covering a larger beneficiary group (beyond immediate project area) and were attributable to the project intervention (contribution analysis). “Normal” project results are reasonably expected and defined in logical frameworks, *Power Results* are *unreasonably expected* and may emerge from conscious effort or simply by chance. A typical Power Result, for example, is national policy change due to pilot testing and innovation.

One Power Result among four others, followed project support to establish 14 effective Fishery Associations to manage the lagoon resource. The key role of these community groups was enshrined in province-wide regulations, and the Lagoon area marked for Fisheries Conservation Zones was increased. Subsequently, the National Assembly passed the new 2017 Law on Fisheries, and Article 10 recognised (for the first time) the right of Fishery Associations to be responsible for the co-management of aquaculture resources and activities nationwide. Those drafting the Law visited the project to see for themselves.

Power results lift the vision of a results-based donor to think wider, longer and more ambitiously about impacts. Consciously articulating them at project design and midterm reviews can increase the chance that they happen. They need a special label because they are more in the form of ultimate impact desires, over which the project has only a marginal influence: we can blame the project for not testing a pilot that was planned, but not because a successful test did not cause a national policy change by end-project.

Thursday, 4 October 2018  
16:30–17:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

PS S3 Poster Session - Strand 3

## P 36 - Measuring Effectiveness in the Experiment of Inclusive Social Security

*M. Kivipelto*<sup>1</sup>

<sup>1</sup> National Institute for Health and Welfare, Social Policy Research, Helsinki, Finland

This poster is presenting the effectiveness evaluation model in the Experiment of Inclusive Social Security. In the experiment, six municipalities are developing and testing new models for inclusive, empowering and participatory social work in Finland. In the experiment, the effectiveness is evaluated using the KEY-measure and nationwide register data. In this poster, the evaluation model of the Experiment of Inclusive Social Security is presented. The KEY is an online measure for determining the effectiveness of social work intervention with adults. The measure will be filled twice with each client of the experiment. The design and development of the KEY measure utilised a realistic evaluation approach with single case-design. In the experiment, the KEY is integrated to the client data management systems. Mobile applications are also used with clients living in sparsely populated areas. The purpose of the use of information technology is that it makes the evaluation more flexible and independent of time and space. The effectiveness is also evaluated by analyzing the register data of social assistance. The register data is gathered from statistical databases of The Social Insurance Institution of Finland (Kela).

Thursday, 4 October 2018  
16:30–17:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PS S3 Poster Session - Strand 3**

## **P 37 - Evaluation Challenges in the Health Field. Lessons Learnt from Evaluating the Cervical Cancer Prevention Program in San Juan, Argentina**

*M.A. Lucero Manzano<sup>1</sup>, M. Audisio<sup>1</sup>*

<sup>1</sup> *Instituto de Investigaciones Socioeconomicas, Facultad de Ciencias Sociales, San Juan, Argentina*

This poster aims to communicate the challenges and opportunities of evaluation in San Juan, Argentina based on the evaluative research done between 2015–2017 analyzing the case of the Cervical Cancer Prevention Program (PPPCC).

The PPPCC was launched in 2009 focusing on the reduction of mortality due to cervix cancer through better diagnosis, an efficient monitoring & treatment system and a data system to analyze the progress and impact of the program.

The study analyzed opportunities, difficulties and challenges faced by the PPPCC to achieve its main objectives between 2009–2017 in the metropolitan area. The final product compiles lessons learnt and recommendations to improve the results of the program.

The methodology included mixed methods considering statistical data (Ministry of Public Health) and interviews to local and national officials of the program, professionals and women using public services. We choose a GAD (Gender and Development) perspective to visualize learnings, outcomes, women and men participation and how it influences the outcomes achieved. This included a reformulation of the DAC's criteria (Relevance, Effectiveness, Efficiency, Impact, Sustainability) and the inclusion of: Coverage, Participation, Appropriation and Empowerment).

The results shows that even though the program aims to prevent cervical cancer, its progress has been limited due to:

- the lack of inclusion of men as women partners with an active role transmitting HPV and with a dominant role in a traditional society.
- Limited resources for diagnosis and treatment but also to increase prevention actions. Professional assistance is also limited.
- Cultural factors: sectorial work within public administration, traditional ideas of sexuality; duplication of activities, lack of communication between professionals and a weak leadership.
- Centralization in two hospitals, making hard to increase coverage in suburban areas, especially consider the context of poverty of many women. It is needed to reinforce territorial actions to be close to communities.
- Weak action about how to prevent the disease.

In this context, the challenges remain in considering evaluations as an opportunity to learn and debate how programs are being implemented. In Argentina, the culture of evaluation is still a pending task as it is considered as a threat or punitive process.

In the health field it is still strong the idea that evaluation is used to determine the value of an effort in quantitative terms rather than a process of social and cultural construction that allows us to better know the reality, to perceive social problems and develop collective proposals to contribute to social welfare. The challenge remain in providing information to reveal processes that allow opening a space for public debate with the power structures involved in decision-making.

It is needed to incorporate evaluation methodologies in the design and execution of policies to strengthen them; building evaluation indicators for periodic reports and results achieved. This requires evaluation trainings for health professionals to provide them with tools to improve the processes of which they are part of.

All these contribute to better communicate the results achieved in order to visibilize the work carried out by the government and the areas with major challenges.

Thursday, 4 October 2018  
16:30–17:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PS S3 Poster Session - Strand 3**

## **P 38 - Developing Partnerships for Communicating Evaluation Findings Using the Site Web: Experiences of Evidence for Development Initiative in Francophone Africa**

*R. Manouore<sup>1</sup>, Z. Njournemi<sup>1</sup>, V. Tuor<sup>2</sup>, B. Togba<sup>3</sup>*

<sup>1</sup> Health Economics & Policy Research and Evaluation for Development Results Group HEREG- Yaoundé-Cameroon, Monitoring and Evaluation and Operational Research, Yaoundé, Cameroon

<sup>2</sup> Evidence For Development E4D Project, Communication, Accra, Ghana

<sup>3</sup> Evidence For Development E4D Project, Capacity Building, Accra, Ghana

In many low and middle income countries, both individual and institutional evaluators face the challenges to find the appropriate easily access channels for communicating their evaluation findings to the global audiences. While the site web becomes more widely accessible to stakeholders, but most individual evaluators and professionals working in the evaluation organizations do not have access to the site web in African countries. This papers presents the partnerships' experiences of the Evidence for Development (E4D) initiative in Francophone Africa with a recipient organization: Health Economics & Policy Research and Evaluation for Development Results Group (HEREG) in Cameroun.

The E4D is a USAID-funded five year project started in 2015 with the aim to strengthen the capacities in M&E and operational research of twelve organizations in six francophone African Countries. The objective of the capacity strengthening program is to increase the technical and managing capacity of indigenous civil society organizations, institutions and networks in M&E and operational research (OR). One of three intermediate results is concerned with research and evaluation findings disseminated locally, regionally and internationally. Overall, the twelve recipient organizations benefited of several theoretical and practical training sessions in M&E, operational research, financial management and communication. Like in other recipient organizations, about 25 professionals from HEREG strengthened their capacities in M&E, OR, communication, site web development and data visualization. HEREG used this partnership to undertake the design and development of its site web where the evaluation findings are disseminated and accessible to the global audiences. HEREG can now develop and disseminate materials communicating results to different target audiences using the site web implemented in partnership with the E4D initiative. In addition, professionals from HEREG individually or in team have strengthened their capacities for conducting all types of evaluations and operational research using up-to-date both combined qualitative and quantitative methods, design and criteria in the framework of internationally recommended standards and ethical norms in Africa context of resilience and action for different stakeholders including vulnerable populations.

Unlike other low and middle income countries, in Africa conducting evaluations for different development programs/projects and public policies is increasing to become usual among individual and institutional evaluators. But the communication, dissemination and sharing of evaluation findings are scarce and very unusual in the Africa context. The E4D initiative tries to address these challenges facing African evaluators starting with few professionals in twelve organizations of six targeted countries.

Thursday, 4 October 2018  
16:30–17:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PS S3 Poster Session - Strand 3**

## **P 39 - Communicating and Using Evaluation Results in Fragile and Conflict-Affected Countries**

*S. Markova*<sup>1</sup>

<sup>1</sup> The World Bank, Public Opinion Research Group, Washington, USA

The World Bank Group (WBG) uses various evaluation techniques for conducting evaluations of its work in fragile and conflict-affected countries. The WBG's Country Opinion Survey (COS) Program systematically measures and tracks the perceptions of the World Bank's clients, partners, and other stakeholders across the globe in client countries. The objective is to improve effectiveness of the World Bank's work in reducing poverty, strengthening economic development, and building more resilient societies, especially in WBG's client countries affected by conflict, violence, and fragility. The WBG's Country Surveys explore perceptions of its work (speed, effectiveness, relevance, etc.), knowledge, and engagement on the ground to improve World Bank results. Survey respondents typically come from national and local governments, multilateral/bilateral agencies, media, academia, the private sector and civil society. The presentation will showcase experience with conducting Country Surveys in fragile and conflict-affected countries, communicating survey data to WBG internal and external stakeholders, using and embedding survey results in the WBG strategic work at the country, regional, and global levels. The presentation will synthesize experience from recent work done by the Public Opinion Research Group in assessing WBG's efforts in 27 fragile and conflict-affected countries, covering 2012–2018 Country Survey data (collected via qualitative and quantitative methods). The presentation will discuss the use and applicability of Country Survey results based on specific examples.

Thursday, 4 October 2018  
16:30–17:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PS S3 Poster Session - Strand 3**

## **P 40 - Digital Platform for Monitoring and Evaluation AcolheSUS Project: A Proposal for Reflection on the Daily Practices of Health Services**

*D. Monteiro Cavalcante<sup>1</sup>, M.B. Botelho<sup>1</sup>, J.R. Cardoso<sup>1</sup>, E.M. Dos Santos<sup>2</sup>, G. Cardoso<sup>2</sup>*

<sup>1</sup> Brazilian Health Ministry, Coordination of the National Humanization Policy, Brasília, Brazil

<sup>2</sup> Oswaldo Cruz Foundation - FIOCRUZ., National School of Public Health - ENSP/FIOCRUZ, Rio de Janeiro - RJ, Brazil

Creating an institutional culture of data collect and use of indicators for monitoring and evaluation in Brazilian public health institutions has undoubtedly been a major challenge. The AcolheSUS project objectives the qualification of care practices, since planning has taken into account all possibilities of using the information produced, not only giving use value, but also creating mechanisms for to collect and for to use of data as an effective tool in decision making, with the purpose of contributing to the improvement of the performance of the groups established for the project, in the fields of activity. The AcolheSUS roject was built in the light of principles, guidelines and provisions of the Brazilian Humanization Policy, from the Brazilian Public Health System – SUS and results from a collaborative implementation and evaluation and developed by a technical cooperation between the Evaluation Laboratory (LASER), of the National School of Public Health (ENSP), of the Oswaldo Cruz Foundation (FIOCRUZ) and the National Coordination of the Brazilian Humanization Policy, Secretariat of Health Care (SHC), of the Brazilian Ministry of Health. The intending is to materialize the Humanization Policy, as a transversal policy of the SUS, inducing reflections about the “doing” and provoking movements to innovate the health care in the SUS. The proposal breaks with hegemonic barriers existing, with the perspective of sharing knowledge and powers, inducing collective construction, welcoming the diversity of knowledges, based on practices of co-responsibility between managers, workers and users, from the perspective of network care. For this purpose, monitoring and evaluation were considered essential in the implementation of the project, being used not only as a thermometer of the current situation, but as a mechanism to trigger actions for improvement, demonstrating opportunities for action, which envisage more qualified processes, based on objective data. Thus, during the design phase of the evaluation of the AcolheSUS project, it was decided to use analysis domains according to the axes of the project, in order to include the information necessary to induce change in practices. This participatory construction process identified the need to create an electronic platform that shares data and actions of this monitoring, oriented to the use of information in order to promote and trigger processes that feedback and allow the executive groups and the national project management to monitor the implementation, identify the needs for improvements, and promote adjustments needed to achieve the intended objectives. Monitoring and evaluation are essential for quality improvement practices, but specifically for health actions and services, although they are valuable tools, they end up being little incorporated into the daily practice of health production, being restricted to the rendering of accounts to the donor agencies and for this reason, this project chose to create a platform that would dialogue with the users and that, in addition to timely monitoring of the project, would also be a tool to record the use of monitoring and influence in monitoring and evaluation of data in the practice of health services.

Thursday, 4 October 2018  
16:30–17:30

STRAND 3: DEVELOPING THE FIELD OF EVALUATION TO PROMOTE RESILIENCE AND ACTION IN CRITICAL TIMES

**PS S3 Poster Session - Strand 3**

**P 42 - Emerging VOPE Led by Young Evaluators**

*N. Salas*<sup>1</sup>

<sup>1</sup> *RedEvalCR Costa Rican VOPE, Coordination Committee, San José, Costa Rica*

The creation of the Evaluation and Monitoring Network of Costa Rica (RedEvalCR) arises and is gestated due to the inexistence in the country of a democratic and articulated evaluation group, in which people with training, practice and interest in issues of evaluation felt represented. That lack of representation reduces the possibilities to growth as a guild, locally with a lack of knowledge of our needs and possibilities for development and integration, and externally, in a fragmentation and inadequate projection in different areas and sectors of development.

Faced with this reality, a group of young people interested in the evaluation, raised the urgency of building a democratic, inclusive and open space that promotes dialogue and a horizontal collective action. This led to the start of a series of open calls in May 2017 for all people who felt called to the subject of evaluation, in order to create a space that would bring together the diversity of interests and knowledge of the people involved in the practice, research, professionalization and dissemination of the discipline of evaluation in Costa Rica.

As a result of a series of participatory workshops and complementary collaborative activities, RedEvalCR was born, as a non-profit organization, non-hierarchical, voluntary and joint work. This Network seeks to support and technically promote, strengthen, disseminate and expand the practice of a national evaluation culture, with the purpose of contribute with the social progress. Now on, RedEvalCR becomes a protagonist and formal technical reference of the practice and institutionalization of a Costa Rican evaluative culture.

This presentation exposes how the RedEvalCR understands the state of the art of the national evaluative task, and how from a collective and democratic construction, it is possible to generate an organizational structure of work, split into commissions, with responsibilities and specific tasks, whose results converged in a second stage, the establishment of an organizational structure and the final management model of RedEvalCR, as a formal non-profit organization, which provides an added value to the development of the Costa Rican evaluation culture.

Thursday, 4 October 2018  
16:30 – 17:30

STRAND 2: RETHINKING EVALUATION METHODS AND METHODOLOGIES

PS S2 Poster Session - Strand 2

## P 43 - Higher Education in Forced Migration: A Comparative Analysis of Scholarship Programmes for Syrian Refugees Implemented by Giz & Daad

D. Lechner<sup>1</sup>, A. Steinhilber<sup>2</sup>

<sup>1</sup> DAAD, S12 - statistics- monitoring and evaluation, Bonn, Germany

<sup>2</sup> Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, New Perspectives through

Academic Education and Training for young Syrians and Jordanians (JOSY), Amman, Jordan  
Higher education in forced migration: A comparative analysis of scholarship programmes for Syrian refugees across countries implemented by GIZ & DAAD.

**Rationale:** The paper addresses two key challenges public institutions are facing when planning and implementing higher education support programmes:

1. Which interventions have proven to be most effective and efficient and how can they be adapted taking into consideration cross-institutional lessons learnt?
2. How can varying institutional approaches rather create synergies than obstacles?

**Background:** Syrian refugees face burdens accessing higher education in their host countries. A successful tool of providing access to higher education for refugees is scholarship awarding. Once enrolled, refugee students experience many difficulties to meet the demands of their university programmes. Difficulties are psychological burdens through past experience of forced migration, direct and indirect discrimination through fellow students or structures of the university system as well as taking over family responsibilities due to loss of the breadwinner. At the same time, scholarship programmes are very demanding, aiming to create Syrian leaders for the post-crisis reconstruction of Syria. Thus, academic expectations and pressure on scholarship holders are high.

**Three interventions under review:** Both GIZ and DAAD currently implement jointly and individually broad programmes tackling this issue:

The JOSY-project provides access to higher education for Syrians and socially-disadvantaged Jordanians through preparatory courses and scholarships for master programmes in Jordan.

JOSY is implemented jointly by GIZ and DAAD.

1. The HOPES-project, financed by the "MADAD fund" and implemented by DAAD in consortium with Nuffic, British Council and Campus France, also provides better access to quality further and higher education opportunities for young Syrians and youth of the host countries Egypt, Iraq, Jordan, Lebanon, and Turkey. Besides scholarships, information on quality BA/MA programmes in the region, counselling services, and language courses are offered.
2. The Leadership for Syria-project provides preparatory courses and scholarships for Syrians who are willing to study in Germany.

**Objectives of meta-analysis:** Firstly, the specific roles and challenges of inter(national) partnerships between public institutions and universities in the context of refugees and higher education are disclosed. Secondly, the effectiveness of the different interventions and their varying modules are analysed in-depth. Modules range from blended learning, counselling concepts and preparatory courses to scholarships across the MENA-region and Germany. Overall, this meta-analysis contributes to cross-institutional learning outcomes for effective intervention-planning.

**Methods:** Within each project, qualitative and quantitative data was collected through surveys, interviews, and focus group discussions from scholarship holders, university teachers and implementers regarding the relevance, efficiency and effectiveness of modules. This data triangulation ensures validity of data. As significant overlap of study design across project-evaluation exists, direct cross-project comparisons were possible.

**Conclusions:** Lessons learnt for future interventions for Syrian refugees show that when designing interventions, various aspects need to be considered: E.g. level of societal pressure, duration which refugees have spent outside their country and their former educational system, the access to and familiarity with blended learning as well as access to a protected study environment with peers. Against this background, success factors for inter-institutional cooperation in the field of higher education are presented.

Thursday, 4 October 2018  
16:30–17:30

STRAND 1: EVALUATION SYSTEMS AND APPROACHES IN TURBULENT TIMES

PS S1 Poster Session - Strand 1

## P 44 - Resilient Schools Embrace Social Innovation: The Case of Family Group Conferences

*L. Zaquini<sup>1</sup>, F. Tarantino<sup>1</sup>*

<sup>1</sup> Associazione per lo Sviluppo della Valutazione e l'Analisi delle Politiche Pubbliche, ASVAPP, Torino, Italy

Is the school system able to adopt innovative interventions in order to face scarcity of welfare resources in a time of increasing welfare needs'?

Family STAR Project aims at preventing social exclusion by reducing school difficulties and dropping out risk for vulnerable students at lower secondary level. Schools were proposed a new intervention model already successfully adopted in the Anglo-Saxons countries in many different fields: Family Group Conferences (FGC) are an operative, light and not expensive tool that aims at detecting students difficulties in a prevention perspective.

Family Star project tested in the Italian lower secondary schools and through a randomized controlled trial this new tool, which involves families – supported by facilitators – with the goal of producing a personalized educational plan for vulnerable students. A FGC is a structured meeting between the members of the “extended family”, one or more teachers and significant others; furthermore, to support the involved student, an external “advocate” is also envisaged. During the meeting, the involved actors discuss the concerns, needs and available resources, in order to identify concrete actions to deal with the student's difficulties and to facilitate the resolution of detected problems. FGC are generally proposed by an institution (for example, by social services or, in this case, by the schools) and organized by a facilitator who does not play a decisional role, but works with the specific purpose of empowering disadvantaged students and their families. The project was implemented in different sites, both urban and rural, in the North and in the South of Italy.

In addition to the randomized controlled trial, the project envisaged an implementation analysis that focused on several aspects: i) school adaptability to innovations; ii) different strategies developed at local level, that strengthen schools resilience capacity; iii) sustainability of the model. The analysis was also guided by the desire to evaluate the challenges faced in conducting a randomized controlled trial of an innovative intervention in the school context.

Specifically, attention was focused on the following phases of the project: training of the new actors required by the model, schools and teachers involvement, students and families participation, facilitators and advocates ability to undertake an empowerment path.

The analysis was conducted during two school years (2016–2017 and 2017–2018). Different data collection tools have been developed: facilitators and advocates online logbooks, satisfaction questionnaires for all the participants in the meetings and two separate questionnaires for teachers. In addition, interviews and focus groups were conducted with school principals, school coordinators, trainers and local project manager.

Lights and shadows of the implementation process were highlighted by the analysis: there are significant obstacles to achieve a full school involvement in the project, due mostly to lack of resources and sceptic attitudes towards innovative models; however, under specific circumstances, evidence shows a general appreciation for participatory and child-centred models.

# Authors Index

Abbot K.	655	Bach H.	527–528, 582	Biřar K.	655
Abma T.	497	Back L.	412	Blackburn J.	379, 421
Abreu D.	248, 739	Baguma E.	355	Blaser C.	216
Aghumian A.	472	Bailly E.	439	Blundo-Canto G.	162, 347, 574–575
Agrawal R.	569	Baker M.	713, 730	Boehmer H. M.	368
Agwu E. A.	198	Balinder Nanda R.	569–570	Bogetic Z.	530–531
Ahmad F.	423	Ballesteros Pena A.	100	Bogetic Ź.	551
Ahonen P.	142	Balzer L.	189	Bohni-Nielsen S.	474
Aibinu A.	330	Bamberger M.	467, 520, 579, 604–605	Bolaños A.	451, 645
Aioanei M.	407	Banerjee P.	193	Bonino F.	194
Akgüngör S.	237, 337, 586	Bangalore Krishnaiah R.	418	Boodhoo A.	220
Al Hudib H.	268, 587	Barajas A.	428, 736	Borges V.	275, 670
Ali S. A.	66	Barbara A. C.	184	Boswell L.	317
Almeida C.	248	Barberi M.	701	Botelho M. B.	721, 737, 746
Almqvist O.	203	Barihahi M.	287	Botzoris G.	218
Alsalem R.	299	Barindelli F.	392	Bourgeois I.	706
Álvarez F.	736	Barreto L.	736	Branco M.	106, 386, 453–454, 509–510, 519–521, 555, 648, 668, 688, 693–694, 699–701, 718
Alvarez P.	210, 322, 525–526	Barreto Lopes L.	720	Bravo-Hernandez J.	527–528
Alves M.	248	Barret S.	628	Brejc M.	84
Amanatidou E.	286	Barrington G. V.	321, 382, 411	Bremner L.	121, 178
Amariles F.	41, 478–480, 501	Baruch B.	509–510, 701	Briton D.	390
Amitsis G.	657	Basile I.	684–685	Brousselle A.	97, 284, 460
Amoatey C.	216	Bastoe P. O.	36	Brousselle D.	329
Ampeh Y.	230	Batista Botelho Alves M.	369	Brown A. M.	641
Anderson S.	628	Bautista E.	231	Bruder M.	626
Andrew P.	169	Bayley J. S.	314, 378	Bruno M.	148, 194
Annan G.	230	Bayley S.	456–457, 464, 682–683	Brüntrup-Seidemann S.	436
Ansell A. E.	260	Baysa A.	711	Buetli D.	706
Antoniazzi R.	343	Beach D.	627	Buhl-Nielsen E.	685
Apgar M.	424, 652–653	Beauchamp E.	94	Bustelo M.	62, 467, 546, 548
Arbulu A.	566	Beaujean S.	352, 396	Bustelo Ruesta M.	660–662, 720
Arévalo J.	239	Bede Benablo G.	77	Butzbach C.	284
Arias Bahia S. C.	369	Belli L.	148	Byeon S. C.	278
Ariss S.	379, 421	Benabaye R.	77, 231	Caballero P.	221
Aronsson I. L.	596, 615	Benaim M.	286	Caballero Partido P.	313
Aronsson L.	617	Bendfeldt-Huthmann L.	217	Caby V.	349
Arranz M.	287, 579–580	Bergstrom M.	164	Čagran S.	82, 84
Arlmann E.	97	Berner P.	234	Callegari S.	681
Asenjo Ruiz C. A.	59, 274	Bertermann K.	417, 544	Campbell D.	439
Atanesyan K.	429	Besrest V.	70	Canares M.	415
Altstrom K.	93, 172	Beuselinc E.	101	Canu E.	222
Atweam D.	230	Bhandari P.	713, 730	Cardoso G.	97, 248, 721, 737, 746
Audisio M.	743	Biesbrouck K.	487–488	Cardoso J.	248
Auriacombe C.	91	Birner R.	153	Cardoso J. R.	746
Awasi R.	183	Bisiaux R.	48, 129	Carneiro G.	129
Ayifah E.	196	Bisson C.	77, 231	Carpinter B. C.	700
Azevedo A.	179				
Azuba R.	293				

Casanova A.	712	Danaj S.	269	Dos Santos E. M.	721, 737, 746
Cavalcante D. M.	721, 737	Daniel E.	149	dos Santos Melo A. A.	369
Cekan J.	620–622	Danilov V.	629	Dotto D.	482
Cekanova J.	464, 607–608	Das A.	214, 250	Doughty W.	147
Celano S.	243, 294	Dasmarinas E.	77, 231	Douthwaite B.	423, 652–653
Çelik Ş. G.	337	Dasmariñas J.	202	Duffy C.	112
César Luiz S.	248	Davidson M. F.	129	Duffy G.	516
Chah K.	265	Davies I.	463–464, 493	Dujso E.	701
Chambel A.	380	Davies R.	640	Dwivedi R.	334
Chambille K.	487–488	de Bruijn M.	686	Ebling G.	38, 49
Chandan A.	713, 730	de Cheveigné S.	62	Echavez C. R.	225
Chandurkar D.	438, 445	de Francisco A.	736	Echavez L. W.	99
Chang C. C.	108, 413, 426	de Goede I.	296	Edwards S.	656–657
Chaplowe S.	65, 579–580, 608, 623	de Jong G.	338	El Khoury de Paula A.	524
Chapman S.	363	de Laat B.	31, 51, 482	Elkins C.	256
Chatsiou K.	78	Delahais T.	103, 135, 420	Elliott C.	261, 474
Chen M.	76–77, 231	de Lancer Julnes P.	95	El Saddik K.	44, 308, 618
Chiba N.	336	Deligiannis A.	201	Emam E.	437
Chikodzore D.	293	De Loughry N.	731	Engebretsen E.	62
Christie C.	69, 134, 474–475, 517	De Luis Iglesias R.	416	Erkan V.	337
Church W.	242	De Maesschalck F.	203	Etta F.	98
Cleaveland A.	65	Demlew T.	733	Fabra-Mata J.	224
Clément L. F.	103	Deng K.	714	Fan C.	157
Cloete F.	91	Denvall V.	271	Fantini L.	182
Colaco R.	77, 231	Deprez S.	582, 613–614, 652–653	Farina S.	255, 281
Colvin J.	236, 287	DErrico S.	43–44	Faure G.	162
Contandriopoulos D.	329	Desmedt E.	154	Faust J.	377
Cook A.	515, 552–553, 558, 563, 644	Desole R.	516	Fazeela H.	647
Copestake J.	288	Devaux A.	420	Febles Carmona E.	188
Coppens V.	160, 262	Devaux-Spatarakis A.	135, 347	Febles E.	320
Corciega J. O.	77	de Villalobos E.	511, 513	Feinstein O.	170
Cordeiro Pereira Cardoso G.	369	Dewachter S.	226	Felcis W.	51, 509–510, 523
Cosentino C.	425	Dewi D. R. A.	362	Félix C.	669
Cousins B.	268, 539, 585, 587	Dey A.	541–543	Fesenko S.	707
Cox D.	286	Dianah M.	729, 738	Figueiró A. C.	739
Crespo C.	75	Dignan T.	223	Fini V.	294
Criels B.	266	Di Gropello E.	472	Fitzpatrick E.	237
Crijns M.	102, 634–636	Dillon B.	280, 612	Flentge J.	558, 564
Cruz M.	709, 712	Dillon N.	537	Flichy A.	70, 135
Culhari T.	386, 521, 718	Dimelu M.	715	Florian M.	187
Cummings H.	411	Dimitrov T.	593	Flueckiger R.	77, 231
Cummings R.	456–457	Dixon V.	348	Foley S.	398
Cuna L.	631, 657	Djapovic A.	567	Forsyth Queen E.	671–672
Dąbrowska-Resiak J.	443	Đokić I.	173	Fowler A.	736
Dahler-Larsen P.	206, 635, 660–662	Dolder O.	406	Franco Alonso S.	100
		Donaldson S.	181	Franco de Abreu D. M.	97
		Donayre Pinedo M.	100	Frankenberger T.	34, 461, 664
				Fred M.	161

Freer G.	556–557	Guinot M.	516	Hurrell A.	499–500
Frölich M.	196	Guillard A.	204	Hutchings C.	34, 664
Froncek B.	110	Gülcan Y.	337	Hyvarinen J.	71
Gaarder M.	348, 669	Gurova G.	361	Hyytinen K.	72, 246
Gadsby E.	185	Gutheil M.	323	Iacob T. D.	660–662
Gaffney S. K.	414	Gutu B.	602	Igarashi M.	148
Gagliardi D.	286	Haarich S.	339	Imata K.	336
Galloway E. T.	568	Hachem S.	205	Inrombe Yano T.	726
Gambardella D.	325	Hainzelin E.	162	Inyang J.	372
Gandhi V.	605, 615	Hall J.	208	Ioannou D.	481–482
Gandin J.	647	Hametner M.	358	Isinika A.	86
Garcia H. J.	148, 293	Hamza C.	341	Isola-Miettinen H.	124
Garcia J.	472	Hamza R.	174	Ison R.	165
Garcia O.	644	Hanberger A.	374, 498	Jackson P.	205
Garduno J.	392	Hanchar A.	613–614	Jacquemin A.	741
Gargani J.	383, 701	Harnar M.	128, 667	Jakoet M.	143, 448
Gault G.	659	Harte E.	316	Jansen van Rensburg M.	502
Gauthier B.	328, 411, 509–510	Harten S.	67	Jarmai K.	385
Gayfer J.	311	Harutyunyan A.	588, 591	Jatoba A.	267, 709
Geeta M.	242	Hasegawa Y.	681	Jean I.	609, 671–672
Georgiadis G.	233	Haslie A.	306	Jeans H.	287
Gerovasili O.	32	Hassall K.	298	Jérôme G.	645
Ghinoi S.	243	Hassannejad O.	244	Johansen K. J.	146
Giffoni F.	441	Hassnain H.	566, 595–596, 615–616	John Michael Maxel O.	80, 155, 227
Gilbreath D.	507	Hawkins P.	471	Johnson K.	472, 632
Gillego F. A. R.	231	Heath S.	261, 344, 474	Jones N.	181
Giovanella L.	712	Heetman A.	139	Jon K.	33
Gobin S.	398	Heider C.	28, 36, 657	Josephine W.	158
Goetghebuer T.	152	Heldeweg M.	362	Julia B.	645
Goetsch E. E.	81, 156, 171, 228, 371	Hense J.	558–564	Julnes G.	539
Gold S.	197	Herforth N.	436	Jung J. D.	716
Gomes N.	352, 386, 396, 521, 718	Herrera F.	130	Kabad J.	739
Gonçalves Marques F.	343	Hilden M.	634–636	Kabell D.	311
Goodman S.	112, 327	Hildenwall V.	432, 679, 696–697	Kabiru D. K.	601
Gorrell P.	184	Hillman J.	128	Kahlert R.	269
Gotsis I.	111	Hirano Y.	150	Kalugampitiya A.	58, 191, 447, 569, 618
Gotz G.	550	Hofman J.	323, 340, 364	Kane R.	326
Gouvalias G.	302	Holvoet N.	226, 266	Kantsos L.	619
Govender J.	183, 242	Homsí M.	301	Kaplan M.	436
Greene J.	381, 498	Honoré V.	70, 135	Karakoula P.	111
Greenwood J.	714	Hooge S.	70	Karasi-Omes C.	740–741
Griffiths S.	140	Hoosen F.	645	Karayannis A.	302
Guertin M.	414	Horvath V.	316	Karijo E.	186
Guffler K.	626	Hotham S.	185	Karlovec K.	666
Guidoccio A.	238, 651, 687–688, 692–694	Houghton L.	244	Karoukis H.	302
Guijt I.	296, 613–614	Huergo E.	428	Katseli L.	32
		Hundt V.	217		

<b>Kaysner K.</b>	59	<b>LaRue M.</b>	719	<b>Madukwe M. C.</b>	198
<b>Kerr G.</b>	485	<b>Lavelle J.</b>	87, 181, 207, 273	<b>Magdalena C.</b>	249
<b>Khaira H.</b>	507, 553, 655	<b>Lawrenz F.</b>	719	<b>Magoola Okalange J.</b>	184
<b>Khan M.</b>	34, 544, 664	<b>Leahy S.</b>	558, 562	<b>Mahapatra T.</b>	214, 250
<b>Khattri N.</b>	472–473	<b>Lech M.</b>	431	<b>Maharana P.</b>	214
<b>Khayyo S.</b>	188	<b>Lechner D.</b>	748	<b>Makleff S.</b>	392
<b>Kiapoka A.</b>	111	<b>Lee L.</b>	121	<b>Maldonado Trujillo C.</b>	466, 468
<b>Kilima F.</b>	86	<b>Lee L. E.</b>	178	<b>Mallol A.</b>	400
<b>Kinda O.</b>	351, 395	<b>Lee S. W.</b>	318	<b>Mankad S.</b>	334
<b>King J.</b>	272, 499–500, 571–573	<b>Leeuw F.</b>	26, 471	<b>Manouore R.</b>	726, 744
<b>Kisbu-Sakarya Y.</b>	717	<b>Lefebvre P.</b>	70, 442	<b>Mansour V.</b>	215
<b>Kishchuk N.</b>	328, 411	<b>Lehouelleur S.</b>	85	<b>Mantouvalou K.</b>	537
<b>Kitsiou A.</b>	111	<b>Lemire S.</b>	69, 134, 475, 517, 556	<b>Mara M.</b>	475
<b>Kivipelfo M.</b>	742	<b>Leppert G.</b>	431	<b>Maranda M.</b>	724
<b>Kliest T.</b>	45	<b>Le Quesne F.</b>	257	<b>Marangoni A. C.</b>	151
<b>Klugman B.</b>	242	<b>Lewis E.</b>	309, 526	<b>Marjanovic S.</b>	270
<b>Knoll B.</b>	62	<b>Ligero J. A.</b>	127, 720	<b>Mark M.</b>	517–518, 541–543
<b>Kocks A.</b>	195	<b>Liger Q.</b>	132	<b>Markova S.</b>	429, 745
<b>Koleros A.</b>	96, 136, 253–254, 365	<b>Lindkvist I.</b>	122	<b>Marra M.</b>	517–518
<b>Kolodkin B.</b>	279, 491, 592	<b>Lindstrom J.</b>	698	<b>Marston C.</b>	392
<b>Korella G.</b>	476–477	<b>Ling T.</b>	270, 353, 680, 701	<b>Martinho J.</b>	192
<b>Korfker F.</b>	490–491, 593, 665–666	<b>Lin Y. C.</b>	413	<b>Martin L.</b>	511
<b>Kotun J.</b>	430	<b>Lira A. R.</b>	721	<b>Martin P.</b>	347
<b>Koudoumakis P.</b>	218	<b>Liu T. Y.</b>	108, 413, 426	<b>Martins de Souza I.</b>	115
<b>Koutoulakis M.</b>	495	<b>Ljungman C.</b>	46	<b>Martins L. G.</b>	369
<b>Koutsomarkos N.</b>	233	<b>Lloyd R.</b>	163, 190, 649	<b>Martinuzzi A.</b>	57, 385
<b>Kramer J. P.</b>	290	<b>Lo A. Y.</b>	413	<b>Marzocchi C.</b>	286
<b>Krasovska O.</b>	360	<b>Loevinsohn M.</b>	391	<b>Massari R.</b>	531
<b>Krisch F.</b>	187	<b>Lomena-Gelis M.</b>	312	<b>Masset E.</b>	354, 722
<b>Krishnan S.</b>	541–543	<b>Lonsdale J.</b>	701	<b>Massey E.</b>	143
<b>Kroll H.</b>	441	<b>Lopes C.</b>	56, 373	<b>Mateo O.</b>	736
<b>Kuboja N.</b>	86	<b>Lord A.</b>	578	<b>Mathe S.</b>	162
<b>Kumar A.</b>	530–531	<b>Lovato C.</b>	88, 255	<b>Mathieu N.</b>	138
<b>Kunihira A.</b>	184	<b>Lozada A. M.</b>	231	<b>Matson C.</b>	232
<b>Kuprii V.</b>	708	<b>Lublinski J.</b>	234	<b>Matsuda Y.</b>	301
<b>Kuştepelii Y.</b>	337	<b>Luca L.</b>	60	<b>Mattila I.</b>	704
<b>Kyewalabye J.</b>	319	<b>Lucero Manzano M. A.</b>	743	<b>McCarty A.</b>	741
<b>Kyriakidou E.</b>	123	<b>Lucks D.</b>	43–44	<b>McConnell J.</b>	257
<b>Lacerda P.</b>	192	<b>Ludden V.</b>	61	<b>McDonald S.</b>	83
<b>LaChenaye J.</b>	119	<b>Lulofs K.</b>	362	<b>McGuire M.</b>	485
<b>Lahey R.</b>	474	<b>Lumino R.</b>	325	<b>McKegg K.</b>	571, 573
<b>Lähteenmaki-Smith K.</b>	167, 634–636	<b>Luna R.</b>	478–480	<b>Mcpherson A.</b>	83
<b>LaMarche M. K.</b>	612	<b>Lundgren H. E.</b>	36	<b>McQuistan C.</b>	609, 624
<b>Lamas Diogo H. R.</b>	718	<b>Luzot A. C.</b>	194	<b>McWhinney D.</b>	554
<b>Landa A.</b>	736	<b>Lyn Higdon G.</b>	424	<b>Meghir R.</b>	656–657
<b>Larsson B.</b>	147	<b>MacPherson N.</b>	459	<b>Meller M.</b>	196
		<b>Maddick K.</b>	731	<b>Mendoza C.</b>	723
				<b>Mergos G.</b>	482

Merritt R.	185	Nakatani M.	336	Omes C.	396
Meyer W.	55, 104, 509–510, 660–662, 691	Nalubega F.	184	Oosi O.	252
Michael R.	366	Nanda R. B.	40–41	Ordóñez M.	239–240
Michailidis G.	285, 630	Nanyanzi M.	137	Orfanidou T.	492
Middernacht Z.	316	Narayanan R.	149, 263, 491, 521	Orth M.	549–550
Mihalfy S.	341	Narchynska T.	588–589, 591	Ould Abdallahi A.	199
Milan D.	101	Nascimento Pereira L.	369	Ousseni M.	389
Millard A.	120, 432, 680	Naudts B.	37, 49	Owen J.	169, 456–457
Mina Rosero L.	211	Ndala L.	402	Oya C.	68
Miranda A.	209	Negre M.	531	Paez K.	105
Mir C.	73	Negroustoueva S.	40	Palenberg M. A.	524
Mishra A.	213, 423, 713, 730	Neuhaus B.	247	Palmisano M. B.	532–533
Mishra P.	713, 730	Neumann L.	435, 704	Panagopoulos I.	302
Mock N.	391	Ngwabi N.	647, 725	Pancotti C.	338, 520
Moeller C.	54	Nieminen M.	246	Parkkonen P.	252
Montague S.	610–611, 674, 676, 678	Nigatu A.	116	Parnia A.	714
Monteiro Cavalcante D.	746	Niikoi C.	230	Pascual U.	236
Monteiro F. D. M.	737	Nilsson J.	38, 49, 536	Passaro Toledo P.	248
Montresor G.	196	Nixon D.	384	Patel P.	149
Montrosse-Moorhead B.	509, 693–694	Njournemi Z.	726, 744	Patterson C.	307
Moore A.	236	Noltze M.	67, 698	Paulmer H.	259
Moraes É.	369	Nordesjo K.	161	Peck L.	698
Morar G.	183	Nyamhuno S.	277, 305, 422, 727	Penaloza Quintero R. E.	211
Moreau K. A.	344	Nyangara F.	444	Pennarz J.	114, 174, 553, 654
Moreira dos Santos E.	97, 248, 369, 739	Oakden J.	571–572	Perez G.	511–512
Moreland L.	355	Obeng-Aduasare Y. F.	568	Perrin B.	456–457, 469–470, 683
Moreno L.	428	Obuku R.	158	Pettersson J.	698
Morier C.	141	Ocampo A.	43–44	Pierobon C.	434
Morrow N.	310	Odhiambo K.	304, 598–599, 729, 738	Pinat J.	148
Mossisa Y.	116	Ofek Y.	53	Piper V.	110
Motta-Moss A.	724	Ofir Z.	35, 459, 540, 578, 620–622	Pirozzi S.	294
Mouret N.	347	Ofori A.	230	Pistis S.	222
Mousa Z.	51	Öhler H.	531	Pittman A.	399
Mrinska O.	630, 685	Ojok J.	333	Pleger L. E.	26
Mukhtar-Landgren D.	161	Ojok R. R.	527–528	Plimakis S.	95
Mukute M.	287	Ojo O.	715	Podems D.	272, 526
Muñoz Muñoz N.	720	Olander S.	697	Polak J. T.	626
Murphy B.	583	Olaolu M. O.	198	Pollard A.	345, 470
Murphy Michalopoulos L.	535	Olavarria C.	40, 238	Poo R.	221
Mustafa G.	74	Olazabal V.	666	Popelier L.	229, 241
Mutambara D.	332	Oliveira A.	335, 709	Porta A.	736
Mwaijande F.	367	Oliveira Cardoso M. R.	369	Porta Lledó A.	720
Mwangi D.	600	Oliveira E.	248	Powell S.	640–641
Naidoo I.	484, 644	Oliveira J.	192	Prasad Pokharel A.	405
Nakaima A.	419, 541–543	Oliveira L. Z.	267, 709	Pratt D.	88
		Oliveros J.	330	Premakanthan S.	264, 594
		Oliver P.	523	Pretari A.	320, 633
				Pritchard R.	324

Proctor M.	533	Rosario C.	369	Schroeter D.	409, 558–564
Protopapas A.	218	Rothman J.	47	Schuetz T.	414
Punton M.	190	Rotondo E.	565–566, 597	Schultz J.	192
Puri J.	34, 303, 577, 664	Routzouni A.	201	Schwandt T.	27, 36, 468, 483, 510, 540
Qaryoufi M.	586	Rowe A.	289, 460	Seabi M. A.	710
Raab M.	276	Roxin H.	195	Sefa E.	230
Raetzell L.	203	Ruiz Villalba A.	644	Segone M.	40, 465, 588, 591, 643–644
Raffree L.	545	Rujiranupong A.	731	Seid A.	733
Rai A.	713, 730	Rusticus S.	88	Seiwald R.	511, 514
Raimondo E.	291, 504–505, 579, 606, 627	Rutu S.	246	Serejo F.	386, 521, 718
Raj A.	250	Ryvudzo N.	183	Serrano Garijo P.	100
Räkköläinen M.	177, 295	Saari E.	72, 634–636	Serre E.	95
Ramasobana M.	350	Sabarre N.	207	Shah G. M.	423
Ramshaw G.	113	Sagara B.	461, 535, 663	Shepherd R.	674
Rasulova K.	588, 590–591	Saha A.	732	Shin J. E.	318
Rauh S.	120, 680	Sai Mala G.	250	Sibanda A.	51, 458–459
Rausmaa S.	252	Sakwa J.	601	Sibanda E.	511, 513
Ray Saraswati L.	213, 713	Salas N.	451, 646, 747	Siblini K.	700
Reda F. H.	176	Salinas S.	51, 258, 458, 478–480, 502	Silva Villanueva P.	64, 632
Reddy S.	56, 309, 525–526	Salminen V.	246	Silvestrini S.	703
Regan E.	77	Salvatori G.	339	Simon B.	104
Reid K.	556–557	Sanchez G.	40	Simon C.	257
Reineck D.	234	Sánchez Romero G.	41, 73, 238, 645	Simons H.	495, 498
Reinhardt G. Y.	78	Sandström U.	62	Singh M.	251
Remr J.	346, 660–662	Sanopoulos A.	341	Singh N.	283
Renmans D.	266	Santiago F.	89	Sinsomboonthong T.	731
Rey L.	375	Santos C.	215	Skalidou D.	68
Reynolds M.	165	Santos M. A.	248	Skiti V.	183
Ribeiro Queiroz de Araújo L.	386, 521, 718	Sanwal N.	334	Sladkova B.	584
Ricardo Lamas Diogo H.	386, 521	Saraswati L. R.	730	Smets L.	531
Richards A.	519–520, 699–700	Sarmento E.	331, 401, 520, 668, 700	Smit C.	143, 200, 404, 448
Richert W.	487–488	Sauvinet-Bedouin R.	153	Smith L.	175
Risby L.	115, 192	Savvakis I.	450	Smith M.	574
Rishko A.	693–694	Sayed Khan R.	511–512	Smith R.	574–575
Rishko-Porcescu A.	509–510, 588, 590, 689	Scanlon K.	657	Smith V.	651, 687
Rist R.	691	Scantlan J.	462, 584	Smits P.	92, 375
Rivera P.	392	Scheers G.	574–575, 582	Snow J.	128
Robert L.	261	Scheinert L.	626	Spanache I.	440
Robinson J.	658	Schmidt-Abbey B.	165	Spilsbury M.	644
Rocha J.	276	Schmidt B.	130	Sridharan S.	419, 541–543, 578, 608, 620–622
Rodriguez-Ariza C.	116, 524, 546	Schmidt J.	504–505	Srikantiah S.	214
Rodriguez-Bilella P.	546	Schmitt J.	550, 625	Stadtmueller H.	546
Roggemann H.	195	Schmitz S.	101	Stame N.	517–518, 538
Rohmann A.	110	Schmuedderich S.	341	Starc N.	173
Rohmer B.	125	Schönherr N.	57	Statman J. M.	260
		Schrader K.	511–512	Stawicki A.	342

<b>Stein D.</b>	254	<b>Tucker S.</b>	272	<b>Werneck N.</b>	192
<b>Steinhilber A.</b>	748	<b>Tuor V.</b>	744	<b>Wheeler P.</b>	232
<b>Stephens A.</b>	309	<b>Tyrrell A.</b>	453–454	<b>Whynot J.</b>	180, 674–675, 678
<b>Stern E.</b>	29, 484, 504–505, 539	<b>Tyrrell T.</b>	469–470	<b>Widmer T.</b>	558–564, 660–662
<b>Stockmann P. D. R.</b>	104, 546, 660–662	<b>Tzavaras Catsambas A.</b>	32, 51, 467	<b>Wilhelmsen A.</b>	122, 224
<b>Stola I.</b>	101	<b>Uitto J.</b>	460, 473	<b>Willacy E.</b>	232
<b>Stoff C.</b>	63	<b>Ulicna D.</b>	659	<b>Williams A.</b>	420
<b>Stryk R.</b>	654–655	<b>Ulla Díez S.</b>	221, 313	<b>Williams B.</b>	546
<b>Stuppert W.</b>	282	<b>Umans H.</b>	207	<b>Wilson D.</b>	623
<b>Suelt Cock V.</b>	211	<b>Uusikylä P.</b>	245, 619	<b>Wilson-Grau R.</b>	292, 527, 574, 581, 638–639
<b>Sumpor M.</b>	173, 630	<b>Uzunkaya M.</b>	331, 666	<b>Wilson K.</b>	394
<b>Sutherland A.</b>	353	<b>Vaca S.</b>	384, 546–547, 642, 735	<b>Wiman A.</b>	385
<b>Sword-Daniels V.</b>	64, 632–633	<b>Vaessen J.</b>	291, 465, 504–506, 530, 690–691	<b>Winckler Andersen O.</b>	685
<b>Sylvestre J.</b>	674, 677	<b>Valant V.</b>	38, 49	<b>Wind T.</b>	168, 315
<b>Tagle L.</b>	243, 294	<b>Valovirta V.</b>	359	<b>Wingender L.</b>	583
<b>Tamondong S.</b>	297, 596, 615–616	<b>van den Berg R. D.</b>	357, 465, 578	<b>Winter K.</b>	126
<b>Tarantino F.</b>	750	<b>Van der Jagt N.</b>	487, 489	<b>Wroblewski A.</b>	133
<b>Tarazona C.</b>	199	<b>van Egmond M.</b>	110	<b>Wyatt A.</b>	384
<b>Tarazona M.</b>	257	<b>van Hoed M.</b>	107	<b>Yaron G.</b>	535
<b>Tarsilla M.</b>	90, 409, 478–480, 546–547, 558	<b>Van Nierop P.</b>	536	<b>Yasin A.</b>	433
<b>Tatarenko A.</b>	477	<b>Van Parijs M.</b>	145	<b>Yaya G.</b>	389
<b>Taube L.</b>	660–662	<b>Vataja K.</b>	252	<b>Yoos A.</b>	232
<b>Taye B.</b>	116	<b>Vazakidis A.</b>	123	<b>Zakariaou N.</b>	389
<b>Taylor-Dormond M.</b>	593	<b>Vedung E.</b>	388	<b>Zaquini L.</b>	750
<b>Taylor J.</b>	701	<b>Vehkasalo V.</b>	219	<b>Zavala I.</b>	392
<b>Tchamba G.</b>	212	<b>Vela C.</b>	118	<b>Zaveri S.</b>	458, 502, 578
<b>Teisen M.</b>	408	<b>Velloo I.</b>	327	<b>Zenker A.</b>	441
<b>Teixeira P. J.</b>	335	<b>Ventimiglia A.</b>	439	<b>Zheenbekov B.</b>	588–589, 591
<b>Teklu A.</b>	733	<b>Verkaart S.</b>	624	<b>Zintl M.</b>	704
<b>Temple L.</b>	162	<b>Verrinder N.</b>	384	<b>Zounkifrou M.</b>	389
<b>Tenev S.</b>	592	<b>Verspohl I.</b>	67	<b>Zwane K.</b>	384
<b>Terra Polanco V.</b>	239–240	<b>Vignetti S.</b>	338		
<b>Thakrar J.</b>	556–557	<b>Villa C.</b>	77		
<b>Theodorakis G.</b>	477	<b>Virginie L.</b>	442		
<b>Tiernan A.</b>	527, 529	<b>Virtanen P.</b>	167		
<b>Till T.</b>	187	<b>Vogel I.</b>	190, 649		
<b>Togba B.</b>	744	<b>Volkan E.</b>	586		
<b>Toillier A.</b>	162	<b>Wadiembe S.</b>	184		
<b>Tómasson T.</b>	129	<b>Wallach S.</b>	235, 649		
<b>Ton G.</b>	96, 652–653	<b>Warne R.</b>	65		
<b>Toptsidou M.</b>	339	<b>Watera J.</b>	159, 449		
<b>Torrighiani C.</b>	403	<b>Watts B.</b>	558–564		
<b>Toulemonde J.</b>	205, 470	<b>Wauters B.</b>	627		
<b>Triomphe B.</b>	162	<b>Wedel R.</b>	195		
<b>Trofin L.</b>	144	<b>Wehipeihana N.</b>	571–572		
<b>Tsuruoka H.</b>	301	<b>Wells M.</b>	236		

