The 9th European Evaluation Society International Conference

6–8 October 2010
Prague, Czech Republic

Evaluation in the Public Interest:
Participation, Politics and Policy

Pre-conference training sessions
4–5 October 2010

Abstract Book

www.ees2010prague.org
Contents

Oral Presentations

5
S5-01 Panel Discussion . . . . The role of country level evaluation capacities

S4-01 Paper Session . . . . Using evaluation in the public policy arena

S2-01 Paper Session . . . . Evaluation within complex political environments

S4-06 Paper Session . . . . Sectoral policy: Labour policy and SMEs

S3-01 Paper Session . . . . Logical frameworks and theories

S3-06 Paper Session . . . . Evaluation use, evaluation impact I

S5-06 Paper Session . . . . Evaluating evaluation

S2-06 Paper Session . . . . Evaluation of public sector policies and programmes

S1-01 Paper Session . . . . Tools and methodologies I

S1-06 Paper Session . . . . Ethical Issues in evaluation practice

S4-11 Paper Session . . . . Sectoral policy research and development

S1-18 Panel Discussion . . . Trends and issues in European university programs in evaluation

S2-07 Paper Session . . . . Evaluating EU Educational policies & systems

S2-02 Paper Session . . . . Evaluation and Policies: Implementation and effects

S3-33 Paper Session . . . . Strengthening evaluation thinking through external evaluation: Rewards and risk

S5-02 Paper Session . . . . Evaluation and capacity development

S3-07 Paper Session . . . . Evaluation use, evaluation impact II

S4-02 Paper Session . . . . Methodology and sectoral policy

S1-02 Paper Session . . . . Methodological challenges I

S4-38 Paper Session . . . . New tools for development evaluation

S4-12 Paper Session . . . . R&D Evaluation: Regional, National, International

S3-12 Round Table Discussion Evaluation Science: A vision for evaluation in public interest

S3-02 Symposium . . . . . . The role of Non-Scientific Knowledge in Evaluation

S1-03 Symposium . . . . . . Developing capability

S4-03 Paper Session . . . . Policy evaluation and rural development

S2-03 Paper Session . . . . Influences on evaluation Processes and Results

S5-03 Paper Session . . . . Counterfactual analysis

S3-08 Paper Session . . . . Performance measurement

S5-08 Paper Session . . . . Evaluation of innovation policies and programmes

S2-08 Paper Session . . . . Evaluating legislation

S3-03 Paper Session . . . . Theories and visions I

S1-08 Paper Session . . . . Partnership and policy


S2-11 Round Table Discussion Creative imagination in interactive evaluation

S1-09 Paper Session . . . . Developing ethical capability

S4-04 Paper Session . . . . School evaluation

S2-04 Paper Session . . . . Interaction between Evaluation Results, Policy and Governance

S4-09 Symposium . . . . . . Sectoral policy evaluation symposium I

S5-04 Paper Session . . . . Towards more rigorous evaluation of development impact

S3-09 Paper Session . . . . Evaluation use, evaluation impact III

S5-09 Panel Discussion . . . Gender and evaluation

S2-09 Panel Discussion . . . Politics of evaluating in and on Violently divided societies

S1-04 Paper Session . . . . Ethics and independence

S3-04 Paper Session . . . . Empowerment
<table>
<thead>
<tr>
<th>Session</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>S4-14</td>
<td>Evaluation and the care sector</td>
<td>72</td>
</tr>
<tr>
<td>S2-12</td>
<td>Round Table Discussion Evaluation System in the public interest</td>
<td>74</td>
</tr>
<tr>
<td>S1-10</td>
<td>Symposium Evaluation competencies</td>
<td>75</td>
</tr>
<tr>
<td>S4-05</td>
<td>Paper Session The Interface of policy and research</td>
<td>76</td>
</tr>
<tr>
<td>S2-05</td>
<td>Paper Session The role and use of evaluations in the policy making process</td>
<td>78</td>
</tr>
<tr>
<td>S4-10</td>
<td>Paper Session Evaluation in childrens and youth services</td>
<td>80</td>
</tr>
<tr>
<td>S5-05</td>
<td>Paper Session Evaluation of natural disasters</td>
<td>82</td>
</tr>
<tr>
<td>S5-23</td>
<td>Panel Discussion Global networking to promote the evaluation profession</td>
<td>83</td>
</tr>
<tr>
<td>S5-10</td>
<td>Paper Session Monitoring and evaluation bottom up and top down</td>
<td>84</td>
</tr>
<tr>
<td>S3-10</td>
<td>Paper Session Theories and visions II</td>
<td>85</td>
</tr>
<tr>
<td>S1-05</td>
<td>Paper Session Methodological processes</td>
<td>87</td>
</tr>
<tr>
<td>S3-05</td>
<td>Paper Session Disentangling complexities</td>
<td>89</td>
</tr>
<tr>
<td>S2-10</td>
<td>Paper Session Civil societies organizations promoting evaluation</td>
<td>91</td>
</tr>
<tr>
<td>S4-16</td>
<td>Paper Session Evaluation in HIV AIDS programmes</td>
<td>92</td>
</tr>
<tr>
<td>S4-27</td>
<td>Paper Session Evaluation and gender</td>
<td>93</td>
</tr>
<tr>
<td>S1-11</td>
<td>Paper Session Transparency</td>
<td>95</td>
</tr>
<tr>
<td>S3-13</td>
<td>Paper Session Self evaluation</td>
<td>96</td>
</tr>
<tr>
<td>S5-11</td>
<td>Paper Session Evaluation of health sector intervention</td>
<td>98</td>
</tr>
<tr>
<td>S4-17</td>
<td>Paper Session Sectoral evaluation: social services evaluation</td>
<td>99</td>
</tr>
<tr>
<td>S3-18</td>
<td>Paper Session Feeding evaluation results back into policy I</td>
<td>101</td>
</tr>
<tr>
<td>S2-13</td>
<td>Paper Session How European Governments treat with evaluation processes and results</td>
<td>103</td>
</tr>
<tr>
<td>S3-24</td>
<td>Symposium Evaluating policies, understanding mechanisms and cumulating evidence</td>
<td>104</td>
</tr>
<tr>
<td>S4-22</td>
<td>Paper Session Evaluation policy: Regional development I</td>
<td>105</td>
</tr>
<tr>
<td>S5-21</td>
<td>Paper Session Participatory evaluation</td>
<td>107</td>
</tr>
<tr>
<td>S5-26</td>
<td>Panel Discussion Evaluation of development as social change</td>
<td>110</td>
</tr>
<tr>
<td>S2-14</td>
<td>Panel Discussion How to capture effects in real world public policies</td>
<td>111</td>
</tr>
<tr>
<td>S1-12</td>
<td>Paper Session Tools and methodologies II</td>
<td>112</td>
</tr>
<tr>
<td>S3-14</td>
<td>Paper Session Evaluation and governance I</td>
<td>114</td>
</tr>
<tr>
<td>S5-12</td>
<td>Panel Discussion Systematic Reviews in International Development</td>
<td>116</td>
</tr>
<tr>
<td>S4-18</td>
<td>Paper Session Sectoral policy area: Health</td>
<td>117</td>
</tr>
<tr>
<td>S4-23</td>
<td>Paper Session Evaluation policy: Regional development II</td>
<td>119</td>
</tr>
<tr>
<td>S5-17</td>
<td>Panel Discussion Sharing lessons of experience</td>
<td>121</td>
</tr>
<tr>
<td>S3-23</td>
<td>Paper Session Participants and co-production I</td>
<td>122</td>
</tr>
<tr>
<td>S3-19</td>
<td>Paper Session Feeding evaluation results back into policy II</td>
<td>124</td>
</tr>
<tr>
<td>S5-22</td>
<td>Paper Session Evaluating the complex</td>
<td>126</td>
</tr>
<tr>
<td>S4-28</td>
<td>Paper Session Evaluation in education</td>
<td>128</td>
</tr>
<tr>
<td>S5-27</td>
<td>Round Table Discussion Evaluation of water related interventions</td>
<td>130</td>
</tr>
<tr>
<td>S2-15</td>
<td>Paper Session Evaluation of systems and findings at the discretion of the policy makers</td>
<td>131</td>
</tr>
<tr>
<td>S1-13</td>
<td>Paper Session Methodological challenges II</td>
<td>132</td>
</tr>
<tr>
<td>S3-16</td>
<td>Paper Session Joint evaluation</td>
<td>134</td>
</tr>
<tr>
<td>S1-20</td>
<td>Panel Discussion How independent are evaluators and evaluations?</td>
<td>136</td>
</tr>
<tr>
<td>S4-20</td>
<td>Paper Session Evaluation, cohesion and integration</td>
<td>137</td>
</tr>
<tr>
<td>S3-20</td>
<td>Paper Session Impact evaluation I</td>
<td>138</td>
</tr>
<tr>
<td>S5-18</td>
<td>Panel Discussion Bridging M and E</td>
<td>139</td>
</tr>
<tr>
<td>S3-25</td>
<td>Paper Session Participants and co-production II</td>
<td>140</td>
</tr>
<tr>
<td>S4-24</td>
<td>Paper Session Evaluation policy: Policy and rural development II</td>
<td>142</td>
</tr>
<tr>
<td>S4-29</td>
<td>Paper Session Evaluation in research and higher education I</td>
<td>144</td>
</tr>
<tr>
<td>S5-28</td>
<td>Panel Discussion Evaluation field building</td>
<td>146</td>
</tr>
<tr>
<td>S2-16</td>
<td>Paper Session How evaluation can influence positive social change</td>
<td>147</td>
</tr>
<tr>
<td>Session</td>
<td>Title</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>S1-14</td>
<td>Paper Session</td>
<td>Methodological challenges III</td>
</tr>
<tr>
<td>S1-19</td>
<td>Panel Discussion</td>
<td>Impact evaluation in the real world: dogma or good practice?</td>
</tr>
<tr>
<td>S5-13</td>
<td>Panel Discussion</td>
<td>Evaluating good governance initiatives</td>
</tr>
<tr>
<td>S4-39</td>
<td>Symposium</td>
<td>Use of evaluation evidence for informed decision making under managing for performance results</td>
</tr>
<tr>
<td>S3-21</td>
<td>Paper Session</td>
<td>Impact evaluation II</td>
</tr>
<tr>
<td>S5-19</td>
<td>Paper Session</td>
<td>Country based and country led evaluation</td>
</tr>
<tr>
<td>S3-26</td>
<td>Panel Discussion</td>
<td>Thinking learning and doing evaluation: A report on local capacity building</td>
</tr>
<tr>
<td>S4-25</td>
<td>Paper Session</td>
<td>Sectoral policy: Gender</td>
</tr>
<tr>
<td>S5-24</td>
<td>Paper Session</td>
<td>Evaluation and transition economies</td>
</tr>
<tr>
<td>S4-30</td>
<td>Paper Session</td>
<td>Evaluation in research and higher education II</td>
</tr>
<tr>
<td>S2-17</td>
<td>Paper Session</td>
<td>The interplay between politics and evaluation: options stakeholders, priorities and decision</td>
</tr>
<tr>
<td>S3-27</td>
<td>Paper Session</td>
<td>Participants and co-production III</td>
</tr>
<tr>
<td>S4-26</td>
<td>Paper Session</td>
<td>Sectoral policy: Urban and regional development</td>
</tr>
<tr>
<td>S5-25</td>
<td>Paper Session</td>
<td>Evaluation and monitoring of interventions in transition economies</td>
</tr>
<tr>
<td>S5-30</td>
<td>Paper Session</td>
<td>Joined up development evaluation</td>
</tr>
<tr>
<td>S3-22</td>
<td>Symposium</td>
<td>Evaluation cultures</td>
</tr>
<tr>
<td>S5-20</td>
<td>Paper Session</td>
<td>Evaluation in context of persistent gender inequalities</td>
</tr>
<tr>
<td>S1-15</td>
<td>Paper Session</td>
<td>Capacity in national context</td>
</tr>
<tr>
<td>S4-31</td>
<td>Paper Session</td>
<td>Evaluation and youth I</td>
</tr>
<tr>
<td>S4-21</td>
<td>Paper Session</td>
<td>Policy and Public Administration</td>
</tr>
<tr>
<td>S5-31</td>
<td>Paper Session</td>
<td>Evaluation conflict and violence</td>
</tr>
<tr>
<td>S4-32</td>
<td>Paper Session</td>
<td>Sectoral policy: Global Health (Technology)</td>
</tr>
<tr>
<td>S4-34</td>
<td>Paper Session</td>
<td>International development</td>
</tr>
<tr>
<td>S4-37</td>
<td>Symposium</td>
<td>Evaluation of anti-tobacco policies</td>
</tr>
<tr>
<td>S2-18</td>
<td>Panel Discussion</td>
<td>Monitoring and evaluating a complex global campaign</td>
</tr>
<tr>
<td>S2-19</td>
<td>Panel Discussion</td>
<td>Contribution to policy and good governance</td>
</tr>
<tr>
<td>S4-36</td>
<td>Paper Session</td>
<td>Evaluation and youth II</td>
</tr>
<tr>
<td>S3-30</td>
<td>Paper Session</td>
<td>Participants and co-production IV</td>
</tr>
<tr>
<td>S3-32</td>
<td>Symposium</td>
<td>The destructive potential of evaluation and how to avoid it</td>
</tr>
<tr>
<td>S5-33</td>
<td>Paper Session</td>
<td>Monitoring an evaluation in developing countries</td>
</tr>
<tr>
<td>S5-35</td>
<td>Paper Session</td>
<td>Evaluating local development</td>
</tr>
<tr>
<td>S5-07</td>
<td>Paper Session</td>
<td>Evaluating agricultural and rural development programmes</td>
</tr>
<tr>
<td>S4-33</td>
<td>Paper Session</td>
<td>Evaluation of management practices</td>
</tr>
</tbody>
</table>

Sponsored by UNIFEM (part of UNWOMEN)

<table>
<thead>
<tr>
<th>Session</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>S4-19</td>
<td>Paper Session</td>
</tr>
<tr>
<td>S1-16</td>
<td>Panel Discussion</td>
</tr>
<tr>
<td>S2-20</td>
<td>Panel Discussion</td>
</tr>
<tr>
<td>S1-17</td>
<td>Paper Session</td>
</tr>
<tr>
<td>S3-31</td>
<td>Paper Session</td>
</tr>
<tr>
<td>S4-07</td>
<td>Paper Session</td>
</tr>
<tr>
<td>S1-07</td>
<td>Paper Session</td>
</tr>
</tbody>
</table>

**Poster session**

**List of Speakers**

**List of Keywords**
The role of country level evaluation capacities

From policies to results. The role of country evaluation capacity development. UNICEF

M. Segone, C. Heider, B. Sanz Luque, M. Saraswathi

1 Senior Evaluation Specialist, UNICEF Evaluation Office, New York, USA
2 Director WFP Evaluation Office, Rome, Italy
3 Head of Evaluation, UNIFEM, Evaluation Office, New York, USA
4 Director and Chair, UNDP Evaluation Office, New York, USA

This symposium is based on a new publication (that will be distributed to participants free of charge) on the role of country evaluation systems to address the policy implementation challenge. In the last decade, several countries have reformed their social policies. However, at times these policy reforms did not deliver the expected results to the target population. It seems that the main challenge resides in implementing policy reforms, rather than in designing and adopting them.

The publication suggests that country evaluation systems can play a central role in producing the relevant evidence to enhance the capacity of national social protection systems to implement policy reforms, by identifying bottlenecks and providing insights into the effectiveness of the strategies used to implement policy reforms, among other things.

Segone introduces the theme describing how country evaluation systems can provide support for addressing the policy implementation challenge. He proposes a conceptual framework for country evaluation capacity development, explaining that it should be context-specific, address supply as well as demand capacities, and be comprehensive by strengthening individual and institutional capacities while enhancing an enabling environment.

Menon argues that there has been greater investment in the development of analytical methods and approaches from the perspective of development cooperation rather than in recognizing fully the potential of evaluation as part of governance within countries. Menon explains that this recognition must be rooted on a fuller understanding of national ownership and capacity in evaluation than is normally the case. If so, this recognition will transform the understanding of evaluation itself.

Heider suggests three key evaluations questions to be addressed in policy evaluations to generate evaluation evidence that is concrete and usable by policy makers and implementers. She also discusses the implications of policy evaluations for the evaluation principles of independence, credibility and utility, as well as the challenges embedded in addressing the three key questions.

Sanz stresses the role of evaluation to achieve gender equality and women’s empowerment. She calls for the need to rethink and enhance the approaches and methods we use, so that evaluations contribute to social transformation processes, including through policy reform implementation. In concluding, she highlights the key parameters of what constitutes the emerging approach of gender equality and human rights responsive evaluation.
Using evaluation in the public policy arena

Can evaluation provide the information needed? Issues of evaluation quality and limited appraisal

A. Bergschmidt, B. Forstner
vTI Federal Research Institute for Rural Areas Forestry and Fisheries, Institute of Farm Economics, Braunschweig, Germany

Assistance for the Modernisation of Farms (AMF) is one of a package of measures of the Rural Development Programme (RDP), the so called “second pillar” of the Common Agricultural Policy of the European Union (EU). With AMF, farmers receive financial support when investing in animal housing, storage, machinery etc. The measure has a number of goals, including enhancement of the competitiveness of the agricultural sector, support for environmentally-sustainable agriculture and the improvement of animal welfare.

Individual EU member states are obliged to evaluate their RDPs according to a common framework of questions and indicators which has been designed by the European Commission. A number of topics for discussion have been triggered by the Common Monitoring and Evaluation Framework (CMEF), and we would like to highlight the following two:

1. Comparative analysis of ex-post evaluation reports of several Member States indicates that the quality of evaluation is often poor as a result of the following factors:
   - lack of adequate data and methodology,
   - use of small-scale surveys of beneficiaries,
   - omission of consideration of the counterfactual situation, and
   - deficiencies in the analysis of causality.

2. Even if these shortcomings in the quality of evaluation could be overcome, it is questionable that information which concentrates on one measure or programme is sufficient for the formulation of well-founded policy recommendations. If there are a number of different policy options for achieving a single goal, the practice of focusing the evaluation on just one of these measures poses a problem.

Options to improve the quality of RDP evaluation include capacity-building among evaluators through the provision of methodological guidelines and an intensified exchange between practical evaluation and the scientific community. To enable evaluation to contribute more effectively to the design of policies, a structured approach to the handling of information on policy alternatives is required.

Keywords: Rural Development Programme; Assistance for the Modernisation of Farms; Capacity building; Evaluation quality

Making the most of Evaluation in Government

S. K. Narayanan, R. Narayanan, P. Gupta, R. Kumar
1 Catalyst Management Services, Monitoring and Evaluation, Bangalore, India
2 Catalyst Management Services, Monitoring and Evaluation, Bhopal, India
3 Catalyst Management Services, Monitoring and Evaluation, New Delhi, India

The paper is derived from recently completed impact evaluation of Agriculture interventions in Tribal Areas in Madhya Pradesh (India). The impact evaluation was commissioned by the State Planning Commission, with the users of evaluation being Agriculture Department and Tribal Welfare Department of the Govt. of Madhya Pradesh. The commissioning authority and users of evaluation outputs were both Govt. bodies. The evaluation scope was vast looking at the multitudes of agriculture programmes and schemes. Such evaluations pose additional issues related to assigning temporal attributability or understanding causality of change.

The impact evaluation involved large scale quantitative study to understand the impact. Absence of baseline and of any relevant progress data posed a huge challenge. The agriculture interventions are spread out across the state and so it was difficult to set a counterfactual as entire population should be equally served.

The recommendations for good practice emerging from this experience are:

- While doing impact evaluation of Govt. programme, look for who has been involved in TOR development, if the respective ministries and departmental views are not strongly taken… limiting the scope of evaluation is not the objective, but generating consensus on the focus of evaluation and its implication for policy is...
- Dissemination need to focus on the mix and a fine balance of micro and macro perspectives that the evaluation throws up. The planners will prefer macro tilt while implementers will prefer micro orientation of the study results
- Post evaluation, there is a need a build a further process of dissemination, testing policy signification of the findings and developing policy notes and finally advocacy for policy changes

The experience shows that impact evaluations of large Govt. programmes have large scope for improving resource allocations, planning and effective targeting, provided the rigour, engagement process and follow up efforts are undertaken.
The institutionalization of evaluation. The “Agencia de evaluacion y calidad de los servicios”: a model for Italy?

V. Andreozzi
University of RomaTre/Uval, Public Policies Economy and Society, Rome, Italy

Programs and policies evaluations are rapidly growing activities raising particular attention within governments of developed countries. In Anglo-Saxon government systems is a practise already implemented, in Italy, public evaluation is not fully covered by scientific studies, nor is it part of technical competences.

Public authorities of continental countries, such as Italy and Spain, are now approaching this issue with a systematic reflection. Since the last period Spanish authorities have begun to be aware of the need to evaluate their actions and in general their ability to generate public value. In 2006 Spanish government created the “Agencia Estatal de Evaluacion de la Calidad de los Servicios y de Politicas Publicas” as an instrument to gradually implement evaluation in all areas of administrative action.

Spain had experimented some evaluation at local level even if in a fragmented way. With the “Agencia”, the Spanish government is trying to constitute a complex and organic system of public evaluation. The goal is to “institutionalize” evaluation, by activating the political process through which the evaluation function is legitimized and developed.

Institutionalization is meant to insert evaluation and its results in the governments agenda and in the development of public policies. Currently in Italy there is much talk about evaluation, referring mainly to a research practice that covers both public programs and policies, as well as organizations and services. What is missing is an organization able to implement both aspects, as it appears to be in Spain with creations of “Agencia”.

The institutionalization of the evaluation will be the main issue of this article, which summarizes particular results of a dissertation work at the University RomaTre I intend to present an analysis of evaluation activities carried out by the Spanish Agency, highlighting points of contact with the institutions, that in Italy, should play a similar role.

Keywords: Sector policy evaluation
Evaluation within complex political environments

Accompanying Research and Impact Analysis of Germany’s Hightech-Strategy – challenges linked with evaluation in a complex policy environment

S. Kind, J. Wessels
VDI/VD E Innovation + Technik GmbH, Socio-Econom ic Section, Berlin, Germany

In August 2006, for the first time ever, the German government has developed a comprehensive national innovation strategy for all its ministries. The aim is to achieve top positions on international research and innovation rankings and to become a top player in the world’s most important markets.

The ministry for education and research developed a concept for an accompanying research of the Hightech-Strategy. Our company has been commissioned by the ministry for a period of three years until January 2011 to support the accompanying research efforts with respect to scientific and organizational aspects. The aim of the accompanying research concept is to monitor and evaluate impacts of R&D activities, initiatives as well as supporting measures and to achieve an overall view of impacts. It supplements foresight activities as well as overarching assessments on innovation processes by two expert groups: The Expert Commission on Research and Innovation and the Industry-Science Research Alliance. Results made available by these different sources target both: to provide input to still proceeding projects, to build the base for conceptualization of new supporting and funding measures and to refine the Hightech-Strategy itself.

The presentation aims to report the special design of the Hightech-Strategy and its implications to the accompanying research concept. It will reflect in what context the accompanying research is embedded and carried out and what kind of methodological challenges are being faced by the evaluators in practice. The presentation shows that the evaluation mainly serves legitimization and communication purposes for political decision-makers. In addition, the presentation focuses on various challenges, e.g. various ministries, divisions and departments are involved; the accompanying research tries to reveal impacts of only recently implemented initiatives and measures. Moreover, the evaluation has to deal with a high complexity due to the high number of programmes, measures, and initiatives being under scrutiny.

Keywords: Accompanying research; Complex policy environment; Legitimization; Hightech Strategy; R&D activities

The politics of ‘outcomes’ in regional program evaluation: arguing over means and ends

B. Taylor¹, R. Eberhard², L. Measson³
¹CSIRO, Ecosystem sustainability, Brisbane, Australia
²Eberhard Consulting, Director, Brisbane, Australia
³AgroParisTech, France

Regional programs aimed at improving natural resource management in Europe and Australia have been criticised by instrumentalists for their seemingly large social and institutional transaction costs at the expense of measurable improvements in water, biodiversity or other benefits. These criticisms overlook the social and political basis underlying resource degradation requiring considerable program investment into the processes of deliberation, problem framing and planning. To counter the argument requires the capacity of evaluators to identify, and participants to articulate, the added value of the regional program in substantive, procedural and politically acceptable terms. This task is far from straightforward however as each is interpreted differently by actors at different scales and again by program designers and implementers. These tensions are heightened when programs being evaluated represent new cooperative modes of governance where the ‘added value’ is characterised by the strengthening of formal and informal ties between community, industry and government actors previously in conflict. We draw on several case examples from Europe and Australia to describe these tensions and discuss the contribution of evaluation in both identifying and articulating procedural, political as well as substantive gains from regional programs.

Keywords: Regional programs; Politics of outcomes; Added value; Scales
United We Stand. Evaluating, Cooperating, and other Unlikely Stories of Evaluation Capacity Building in Italy

L. Tagle, M. Bolli, A. Monteleone, S. Ciampi

1 Inea, Inea Roma, Italy
2 Struttura nazionale di valutazione FSE, Isfol Roma, Italy

After utilizing European Structural Funds evaluation requirements to introduce evaluation practices, Italian authorities are now striving to build national capacity. This effort revolves around a number of innovations:

- a focus on single interventions or coherent groups of interventions – as opposed to whole programmes – regardless of the sources of funding
- the possibility of conducting evaluations internally or externally
- programming and managing authorities’ greater and more explicit responsibility and involvement in evaluation
- the possibility of conducting ex post evaluations.

Together with the weight of previous practices, and in the specific context of the cohesion policy, these innovations create a complex, and sometimes conflict-inducing, framework: how to combine ownership with independence, internal evaluation with credibility, the time-frame of policy with the time and resources required to achieve solid evaluation results? In order to support the innovations, and tackle these issues, the Italian national strategy framework 2007–2013 has reinforced the Italian National Evaluation System, a joint endeavor in which the three central evaluation agencies (the Evaluation Unit of the Ministry for Economic Development and the two agencies respectively responsible for evaluations of rural development and labor policies) and regional Evaluation Units provide guidance and support evaluation processes in Regions and ministries. We will analyze the main activities, challenges, and results of our work towards place-based, effect-focused evaluations. We will particularly address the issue of why, how and to which extent we try to overcome the institutional silos that are characteristic of cohesion policy in Italy at all levels, from municipalities to regions to ministries to European Commission Directorates General.

Keywords: Evaluation Capacity Building; Internal-external evaluation; Institutional silos
Sectoral policy: Labour policy and SMEs

Do non-refundable subsidies promote SMEs growth?

S. Csengodi  
AGENDA Consulting Ltd, evaluation, Budapest, Hungary

We evaluated a measure providing a maximum of EUR100,000 and 50 % non-refundable subsidies for Hungarian SME's development projects [purchase of equipments, licence, know-how and (re)construction of their production buildings]. The Managing Authority decided to subsidize 2530 SMEs of the 7076 applicants between 2004 and 2006.

We measured the – intended and unintended – results achieved by the intervention. We applied a large panel dataset [for the years 2002–2007] of all – subsidized and non-subsidized – Hungarian SME's that allowed us to create a well-defined control group for the subsidized firms. We ran OLS and fixed-effect estimations and applied matching techniques to measure the effects of non-refundable subsidies on the growth of subsidized SMEs (we considered growth of 1.) investments, 2.) turnover, 3.) value added 4.) long-term liabilities).

We find that – as a result of subsidization – non-refundable subsidies have increased the extent of the subsidized SME's investments. We see a slight growth in the subsidized SMEs’ employment. We do not find any other real positive impacts of the subsidies.

Keywords: Non-refundable subsidy; Erdf; Counterfactual; SME; Growth

The effectiveness of Direct Investment Supports for Small and Medium Enterprises

A. Beres  
National Development Agency, Coordination Managing Authority, Budapest, Hungary

The focus of this study is measuring the effectiveness of the direct investment support scheme implemented in the framework of the Economic Competitiveness Operational Programme 2004–06 (ECOP) in Hungary. ECOP was co-financed by the European Regional Development Fund, and offered a wide variety of grant schemes for enterprises, governmental organizations and municipalities in order to strengthen the economic competitiveness of Hungary. The study focuses solely on one component of the program, namely the direct investment support scheme for small and medium enterprises (SME) for technological modernization. It uses micro level data and counterfactual method (propensity score matching) to create a control group and to estimate the direct effects of the intervention.

To summarize our findings on the effects, the support scheme in focus created the opportunity for the assisted enterprises to use additional resources for their operation (both for capital and labor), but these additional inputs did not create extra growth for the enterprises or additional tax revenues for the public budget (compared to the control group). The study also addresses the issue of the so called dead weight, “the change observed among direct beneficiaries following the public intervention, or reported by direct addressees as a consequence of the public intervention, that would have occurred, even without the intervention” (European Commission).

Keywords: Grants for SMEs; Counterfactual method; Propensity score matching; Difference in differences; Cohesion Policy
The evaluation of active labour policies
L. Conzimu, G. Giorgio
Regione autonoma della sardegna, Nucleo di valutazione e verifica degli investimenti pubblici, Cagliari, Italy

Background
The purpose of this study is to compare and analyse the potential gain of integrating two sources of dataset: data coming from administrative registers for the study of the labour market and the information available through field research data. The information extracted from existing administrative registers, at regional and national level, can help to create and promote studies and interpretations of the evolution of the labour market, reducing time and cost of the research activities. On the other hand, surveys should integrate information not available from existing data.

In Italy, the reform of the labour market (Law 59/97, Decree 469/97 and further legislation) moved from focusing on the social protection systems, to a new approach to the problems of unemployment, based on active labour policies, such as actions undertaken by institutions to promote employability. These new laws led to a decentralization of the management system of the labour market at a local level, while at the national level still remains a general role for guidance, promotion and coordination.

Objectives
The Autonomous Region of Sardinia has invested significant resources in recent years in active labour policies, especially through Structural Funds (about 33 M € axis II POR ESF 2007–13), for the implementation of employment services, known as Service Centres for Labour (Centri Servizi per il Lavoro). The research aims at providing useful information to policy makers, both in terms of governance of the CSL and in terms of the effectiveness of services for a better use of resources.

Methods
The tools used are a combination of qualitative and quantitative methods. First we carry out a survey on CSL customers’ with a twofold purpose: customer satisfaction and placement analysis.

On the other hand, data of administrative registers will be used for causal inference, in order to estimate a counterfactual situation.

Results
The evaluation research provides, as a main result, the formulation of a proposal for improving the management of employment services and its implementation.

Keywords: Labour policies; Customer Satisfaction; Placement survey; Employment Service; Administrative registers
Logical frameworks and theories

Theory-based Evaluation of Emergent Programs – a Constructivist Approach
K. Haubrich
German Youth Institute, Youth and Youth Welfare, München, Germany

How is it possible to evaluate developmental multi-site programs with given the presence, and heterogeneity, of several relatively autonomous projects developing and testing innovative concepts in social work? The evaluation of model programmes in the field of child and youth services in Germany has to answer the question whether and how innovative concepts “work” and “produce” the desired outcomes. Program theory offers a way making explicit the assumptions of social workers underlying an intervention and innovative strategies. However, the methodological approaches for reconstructing program theory discussed in the evaluation literature are often rather “vague”. In the evaluation of model programs at the German Youth Institute an innovative methodological approach of program theory evaluation was developed. First, the main characteristics of developmental programs as “evaluands” are described, with a number of model programs in the field of child and youth welfare serving as an example. Second, a reconstructive approach to program theory evaluation is proposed. It provides a concept of theory-based evaluation that allows for an empirically well-grounded theoretical description of a multi-site program that is conceived of as a “cluster” of heterogeneous projects testing similar pedagogical strategies in different contexts.

Contribution analysis: The promising new approach to causal claims
S. Lemire
Ramball Management, Education, Copenhagen, Denmark

Impact evaluation has for many years received continued attention and motivated heated debates within the international development community. Yet, despite the sustained attention to determine the attribution of projects, programs, and policies, there is to this day few methodologically rigorous alternatives to the traditional counterfactual impact designs. Moreover, the received wisdom among many evaluators is that real impact evaluation cannot be done in the absence of controlled comparisons. One promising alternative to the traditional counterfactual designs is that of John Mayne’s Contribution Analysis (CA) that seeks to examine attribution through contribution (Mayne, 1999, 2001, and 2008).

In this paper, I first examine the methodological strengths and weaknesses of CA, awarding particular attention to the ability of CA to identify and determine the extent of influence from alternative explanations. Informed by this examination, I argue that CA – in its current form and application – fares well in terms of identifying external factors, but finds itself in need of further methodological elaboration to adequately account for the extent of influence from these factors. As such, CA remains vulnerable to threats of internal validity. In an attempt to further advance CA, I propose that the sound use of CA demands the consistent and rigorous application of a set of practical techniques that aims to illuminate the influence of external factors.

The overarching purpose of this paper is two-fold, in that I seek to both (1) advance the theoretical discussion on the internal validity of CA and (2) contribute to the soundness of the practical implementation of CA. Indeed, the prospect of strengthening these two aspects of CA is what motivates my effort. The paper will draw on practical examples and experiences from an evaluation employing CA in the field of education.

Keywords: Impact evaluation; Contribution analysis; Methodology; Validity; Education
Assessing the validity of program theory: logic analysis

A. Brousselle
University of Sherbrooke, Community Health dep. – Centre de recherche de l’hôpital Charles-LeMoyne, Montreal, Canada

Program theory evaluation, which has grown in use over the past 10 years, assesses whether a program is designed in such a way that it can achieve its intended outcomes. We describe here a particular type of program theory evaluation—logic analysis—that allows us to test the plausibility of a program’s theory using scientific knowledge. Nowadays, evaluators build logic models to conceptualize interventions based on stakeholders’ perceptions of how the intervention works. Nevertheless, very few evaluators test the plausibility of a program’s theory before using it. Yet, using a program theory that is incomplete or inaccurate can lead to flawed evaluations and recommendations. As Suchman and Weiss stated, the failure to find program effects could, when not attributable to faulty evaluation design, be due either to the wrong theory or to inadequate implementation (Bickman, 1987-b; Birckmayer & Weiss, 2000; Chen, 2004; Weiss, 2007)

In this presentation we will first briefly present the roots of program theory evaluation. We will then define and describe the two types of logic analysis (direct and reverse logic analysis) and present methods for testing the plausibility of a program’s theory. Then we will illustrate the two types of logic analysis with two real evaluation cases, the logic analysis of practice guidelines for patients with mental and substance use disorders and the logic analysis of the Research Collective on Primary Care.

Logic analysis is useful for improving the intervention or finding alternatives for achieving intended outcomes; it influences the choice of type of evaluation to conduct and strengthens the validity of subsequent evaluations. It is an important contribution to program theory evaluation. This presentation will help evaluators in designing and carrying out such projects.

Keywords: Program theory evaluation; Logic analysis; Theory-driven evaluation

Contribution Analysis: A Navigational Tool for Local Partnerships Steering towards Outcomes in Complex Systems

E. Wimbush¹, T. Mulherin²
¹University of Edinburgh, Business School, Edinburgh, United Kingdom
²NHS Health Scotland, Policy Evaluation, Edinburgh, United Kingdom

There is growing discussion in the academic and policy literature about complexity and complex systems as a way of framing the problems posed by a global, inter-connected world. To tackle these problems, new horizontal, inter-organisational forms are emerging (collaborative networks, partnerships) and the traditional tools of control (performance management, budgets) are being superseded by influence and enabling strategies and steering forms of network governance. Compared to the literature on the challenges of results-based management, there has been much less attention given to understanding the implications for evaluation and performance measurement especially in the context of partnership or network governance. If complex, dynamical systems generate non-linear and unpredictable results, a key question for evaluation is how to address attribution (cause and effect) in non-linear systems. Is it possible to delineate the contribution of a specific programme, policy or organisation to a pre-determined ‘shared’ outcome within a complex system? While action research and theory-based evaluation offer broad methodologies relevant to systems change, Contribution Analysis (John Mayne 1999, 2010) and Results Planning (Steve Montague www.pmn.net) offer a tangible, systematic framework for applying these general principles. Structured Contribution Analysis helps to chart and assess a program’s intended process of change and thereby construct and evidence a performance story about its ‘contribution’ to higher level outcomes. This can be adjusted and strengthened over time, taking into account external influencing factors related to its context. However, there are only a few examples of Contribution Analysis being applied in practice, and none so far that have applied this approach within the context of partnership performance. The purpose of this paper is report the findings from current work in Scotland to assess whether and how Contribution Analysis might be applied to local partnership planning, evaluation and performance reporting. The paper reports on two case studies of Contribution Analysis applied to local partnerships in Scotland: one is seeking to address inequalities in mental well-being in a defined community and the other is developing a strategy for tackling anti-social behaviour.

Keywords: Contribution Analysis; Partnerships; Complex systems; Scotland UK
Evaluation use, evaluation impact I

Evaluating the evaluation or from development of theory to advancing evaluation utilization

A. Neuman1, N. Shahor1, A. Sarid1, I. Shina2, Z. Saar2, R. Savaya3
1 Western Galilee Academic College, Education, Kfar Veredim, Israel
2 Karev Educational Program, Evaluation, Shorashim, Israel
3 Tel-Aviv University, Tel-Aviv, Israel

This article presents findings from a two-stage study, carried out by the internal Evaluation Unit of the Karev Educational Program, which examines the way organization members use dozens of evaluation studies performed by the Evaluation Unit.

The study’s first stage addresses development of the theory behind Evaluation Utilization in an educational organization, while the second stage addresses the application of this theory in order to advance evaluation utilization within the organization.

This theory was built using the participatory method of concept mapping, developed and described by Trochim (1989). This process, which was described in depth in Shahor et al (2008), resulted in a map of Evaluation Utilization concepts divided into 11 categories, and generated a theory about the way Evaluation Utilization is used within an organization.

To further the use of Evaluation Utilization, a review of “actual and preferable evaluation utilization in the organization” was performed, using the 11 categories identified in the previous stage.

The review results were also incorporated into the map, providing another dimension – height – that helps understand the correlation between the actual and preferred use of Evaluation Utilization in the organization. The planning of future evaluations will be based on analysis of this correlation.

The results of this study also provide the Evaluation Unit with a framework for: setting targets; compiling an annual Evaluation program; decisions about different evaluation subjects based on their potential application; reflective processes within the Evaluation Unit; a model that other units in the organization can apply to their own work assessment processes; and most importantly, demonstrating a key evaluation principle for the organization: how an Evaluation Unit conducts its decision-making process based on applicable data.

Keywords: Evaluation utilization; Concept mapping; Decision-making process; Evaluating the evaluation; Methodology; Standards; Impact

Agile evaluation of technology and innovation programmes – From auditing to strategic development and utilisation

P. Pesonen
Tekes, Impact Analysis, Helsinki, Finland

The paper studies the evaluation processes in the field of technology and innovation programmes. It is based on the experience gained in Tekes, the Finnish funding agency for technology and innovation. The methodology applied contains literature analysis, observation methods and interviews.

Tekes has carried out more than 200 technology and innovation programmes and commissioned more than 100 evaluations during the last 25 years. There has been several evolution phases of Tekes programmes. The first programmes were small and technology-oriented and their governance was quite simple. The latest programmes are large and complex. At the same time, the evaluation approaches and procedures have evolved. The analysis focuses on modelling dynamics of the evaluation process. We show how evaluation has transformed from auditing perspective towards applied research and then towards future-oriented strategic development.

The authors recommend the need for both applied research and consultancy in evaluations. The practical application is better understanding of the interplay and different roles of commissioner, evaluators and utilisers of results.

Keywords: Technology and innovation programmes; Evaluation process; Commissioner of evaluations
Impact of evaluations matters

D. Svoboda
Svoboda, Prague, Czech Republic

Background
There are never-ending discussions about sophisticated evaluation methods, competencies and independence. In the fight for perfect results, the use of evaluation results is still underestimated. The practice shows significant gaps:

- The decision makers often have no idea how (and why) to use evaluation results;
- The results of development interventions are assessed mainly by donors; and the final beneficiaries have little chance to say what their real priorities are;
- The evaluators do not accept any responsibility for the intervention to be evaluated;
- There are problems in communicating evaluation findings.

Influential evaluations are considered for a specific category and final beneficiaries do not participate in setting and evaluating the real impacts.

Objective
No benefits are sustainable without significant changes in conditions, knowledge, skills, attitudes and behavior. Focus on real benefits must lead to shift from accountability to donors to accountability for impacts on people’s lives.

Methods
Participatory approaches do not mean using the participants as objects for formal approvals of donor strategies, or as objects of evaluators? focus groups, questionnaires or counterfactual experiments. Similarly, not the attribution to donor programs but contribution to sustainable benefits is the core of impact evaluations. The beneficiaries must thus participate in setting priorities of development policies, programs and projects, in setting indicators of benefits that are to improve their lives, and they must enforce their own evaluation questions, appropriate evaluation methods, and utilization of lessons learned. And evaluators must accept their responsibility for impacts of their evaluations and not only for activities and outputs.

Conclusions
There is no partnership without joint accountability and there is no chance to reach sustainable benefits or utilize evaluation results without true ownership and inclusiveness both of development interventions and their evaluations. There is no sense to carry out non-influential evaluations…

References:
• Accra Agenda for Action (2008). §15 “We will strengthen and use developing country systems to the maximum extent possible”; Accra, Ghana, September 2008
• Current EC/EuropeAid Calls for Proposals
• Segone, M., et al. (2009). Country-led monitoring and evaluation systems – Better evidence, better policies, better development results; UNICEF Regional Office for CEE/CIS, Switzerland, 2009
• Tarsila, M. (2010). Being Blind in a World of Multiple Perspectives: The Evaluator’s Dilemma Between the Hope of Becoming a Team Player and the Fear of Becoming a Critical Friend with No Friends; Journal of MultiDisciplinary Evaluation, Volume 6, Number 13, ISSN 1556-8180, February 2010
• ToR for evaluator of the Czech ODA projects under the recent Requests for Proposals
Evaluating evaluation

Evaluation and its institutionalization in Mexico Lessons learned and challenges ahead
G. Perez Yarahuan
Universidad Iberoamericana, Social and Political Science, Mexico. D. F, Mexico

During the past decade Mexico has witnessed important changes on its political system and its policy making process. One of these is the evaluation system for social programs put in place since the end of the 1990’s with the launching of the first conditional cash transfer program (CCT) in Latin America, PROGRESA, and its world wide known impact evaluations. Since 2001, the Mexican government has put an evaluation requirement of its more than 100 social programs. However little is known of how this effort has had an effect on the government’s interventions and policies.

The objective of this paper is first to describe why and how the Mexican evaluation system of social programs has been put into place by this country’s government. What types of evaluation studies are made, what were their requirements on methods and utilization criteria. The second objective is to analyze a group of program s for which evaluations were made during this period and see if any effect can be found regarding its utilization to enhance program performance.

This paper uses a sample of programs to analyze if the evaluations made between 2001 and 2009 of social program s in Mexico have been coherent with program’s changes with regards to formal operation rules. Results of this analysis show that the characteristics of programs, such as budget, time of operation and sector, are variables that affect changes due to evaluations.

This study provides information and evidence on the creation of a government’s evaluation system and the lessons and challenges on the use of evaluation research for policy change in a developing country.

Keywords: Mexico; Evaluation system; Utilization; Institutionalization

Evaluating Evaluation Systems. The German Development Cooperation Case
R. Stockmann, W. Meyer, J. Rech
Saarland University, CEval, Saarbrücken, Germany

Usually the objectives of Evaluations are limited on some specific issues, strategies, projects or programs. In some cases, meta-evaluations are launched for assuring quality standards of a series of evaluations. An analysis of the whole evaluation system, implemented to support policy makers, is still a rare exception.

However, one can find some of these exceptions in the development cooperation sector. Guided by the Development Assistance Committee (DAC) of the OECD, the donor countries try to harmonize not only aid and cooperation between North and South but also the evaluation of development cooperation programs and projects. Nowadays, the DAC criteria and the DAC standards are commonly known among those organisations which are involved in such kind of programs and projects, at least in the OECD-countries. Moreover, the DAC also launches enquiries on the development cooperation systems in donor countries, including some comments on the evaluation practice as well.

While the DAC only uses peer reviews as method for this, the German Ministry of Development Cooperation (BMZ) commissioned a multi-method system check of about 30 different implementation agencies supported by the Ministry. The results of this study are used here as an example for evaluating evaluation systems and for to show the additional surplus of such kind of evaluations. It will be argued, that one has to take the policy sector as a holistic system and evaluation culture is a straight outcome of this system – and less related to the organisational traditions, national evaluation traditions or international scientific debates on evaluation theory.

Keywords: Development Cooperation; Evaluation System; Evaluation Methods
Evaluating the capacity development practices of Belgian development NGO’s: constructing an emerging practice

N. Molenaers¹, G. Phlix², J. Bossuyt¹, B. Fonteneau¹, H. Huyse⁴
¹University of Antwerp, Instituut voor Ontwikkelingsbeleid, Antwerpen, Belgium
²ACE Europe, Management, Mechelen, Belgium
³ECDPM, Development, Maastricht, Belgium
⁴University of Leuven (HIVa), Sustainable Development, Leuven, Belgium

Capacity development support of southern partners has become a cornerstone of the work of many Northern development NGO’s. While there is a growing consensus in the aid sector on the importance of capacity development and some of the basic ingredients of successful external CD support are emerging, little systematic research has been done on the practices of development NGO’s in this area. In 2008 the evaluation department of the Belgian Ministry of Foreign Affairs and Development Cooperation commissioned a large scale evaluation focusing on capacity development support within NGO partnerships. Over the whole of 2009 a consortium of international and national consultants evaluated 31 partnerships of 21 Belgian NGO’s in six countries. Important methodological challenges related to (1) the large diversity in actors, projects, themes and contexts that had to be assessed, (2) the dynamic nature and non-linearity of capacity development processes, (3) and the implicit character of the NGO practice in this area resulting in a lack of secondary data. To deal with these challenges, the evaluation was inspired by and provided a validation opportunity of a number of recent theoretical frameworks, such as the 5 core-capabilities framework of ECDPM (2008) and frameworks of INTRAC (2007, 2010). The methodology made use of conventional data-collection techniques (document study, interviews, surveys, focus groups) in combination with self-assessment workshops in groups, and the collection and analysis of stories. This was done with a view of first re-constructing the emerging practice, and secondly testing the underlying assumptions and intervention logic of that practice. This research paper will describe the main findings of the evaluation, and discuss the experiences with the theoretical and methodological frameworks used, and will highlight emerging issues for these types of evaluations.

Keywords: Evaluation; Capacity development; Development aid; Non governmental organisations; Partnership
Evaluation of public sector policies and programmes

Evaluating empowerment programmes and outsourcing of public services in an Italian city

B. Buffoni
Lattanzio e Associati, Consulting and research, Milano, Italy

Parma is a city in the centre of the Po Valley, with 200,000 inhabitants. It is a city with a strong sense of identity, a rich history – as one of the Italian “small capital cities” between 1500 and 1860, a high level of public services and civic awareness.

Some years ago, after a change in the administration, some innovative policies have been started by the new government. This paper considers the outsourcing of public services and the personnel policy of the administration, and the evaluation that was conducted of these two policies during 2009.

Some public services, traditionally managed directly by the municipality, were outsourced, selecting suppliers through public tender: local transport, waste management, cemetery services and others.

An internal communication and training programme for all employees was started in 2008. The aim was to improve the understanding of the administration policies, and enhance motivation and involvement. It involved 1,237 employees (90% of the total), and was carried out using innovative and integrated methodologies.

Two separate evaluations were conducted: the quality of services was appraised with a customer satisfaction survey carried out in two steps, in order to allow and to measure improvements of the services. It was supported by an audit on the legal implementation of the outsourcing contracts.

The evaluation of the training programme aimed at assessing the impact of the programme on the organizational climate and on the performance of the employees.

The aim of this paper is to illustrate the results of the interventions, the methodologies used for the evaluation, the evaluation results, in terms of improved awareness of the administration and of indications for future activities.

The integrated and coordinated approach to the two policies will be underlined, as well as the relevance of effective empowerment programmes in situations of change.

Keywords: Outsourcing; Empowerment; Customer-satisfaction; Motivation; Local-government

Challenges in evaluation of EU programme and policies, example of linking inputs-activities to effects-impacts in an agency evaluation

M. Rune Jensen
COWI Belgium

The interim evaluation of EACEA (Education, Audiovisual and Culture Executive Agency) was the first evaluation of the Agency’s management of and implementation of the following programmes: LifeLong Learning, Youth in Action, Culture 2007, Europe for Citizen and Media 2007. The evaluation was launched by DG EAC.

The Agency became operational in 2006 and the main tasks of the Agency is to manage projects including launch and conclude grant procedures, call for proposals, evaluation of projects, selection, notification, signing agreements, monitoring and evaluation of projects, payments and contacts with beneficiaries.

The evaluation was carried out in answering 13 specific evaluation questions addressing relevance, effectiveness and efficiency in accordance with the EU DG Budget guideline for evaluation. The evaluation focussed on performance management in the implementation of EU programmes, looking at achievements and impacts, workload indicators and a CBA. The evaluation took into account both the European Commission as well as the beneficiaries’ side.

Main findings:
- The services provided to the beneficiaries and stakeholders by the agency have improved compared to the quality of the services provided by previous management schemes.
- The CBA shows that the management in the Agency is cost-effective and approximately 20% cheaper than the Commission.
- The Agency functions in accordance with its legal framework and there is no overlap of tasks between the Commission and the Agency.
- The Agency has achieved the objectives set in the 2006 and 2007 work programmes as well as succeeding in simplifying its internal procedures.
- The flat organisation of the Agency and the problem solving attitude in the Agency have had a positive impact leading to increased efficiency.

The presentation will focus on the challenges in evaluation of EU programme and policies, in this case lessons learnt from evaluating EU agencies linking inputs-activities to effects-impacts.

Keywords: Performance; Evaluation; EU; Agency; Impacts
Valuing public policies and programs: developing the valuative ground for evaluation in the public interest

G. Julnes
University of Baltimore, College of Public Affairs, Baltimore, USA

Several years ago the U.S. evaluation community became embroiled in controversy over the appropriateness of random assignment experimental methods in establishing causal claims regarding the impacts of innovative approaches. The proximal flashpoint for the controversy was a policy in a federal agency giving priority to these methods, but a deeper cause was a tacit “agreement to disagree” among evaluators from different perspectives that worked against developing any working consensus on the contextual value of alternative evaluation methods.

Considerable progress has since been made in identifying areas of consensus on methods for estimating program impacts, but a parallel challenge is emerging for the methods we use to support judgments of the value of these programs and policies. Most realize that positioning evaluation as something that supports the public interest requires a stance on just what the public interest consists of and how we support value judgments relevant to this interest. However, as government agencies increasingly emphasizing particular approaches to valuation (such as economic valuation), the evaluation community “agrees to disagree” on the value of competing approaches to valuing.

To advance the needed working consensus, a volume of New Directions for Evaluation has been proposed that draws together the insights of major evaluation scholars (mostly U.S. but with some European representation) in three areas: (1) highlighting foundational issues in alternative approaches to judging the value of programs, (2) considering the strengths and weaknesses of different approaches to valuation in different contexts, and (3) developing recommendations for policies that support effective, context-sensitive public-sector valuation. This proposed presentation would provide a brief background on this volume and present a framework organizing areas of consensus on the contextual appropriateness of available approaches to valuing and on valuation policies that can better support evaluation in the public interest.

Keywords: Valuation
Tools and methodologies I

Practical Economic Evaluation Tools for the Social Welfare Field

A. Kettunen, M. Pulliainen
Diaconia University of Applied Sciences, Diak East Research Centre for Social Economics, Pieksämäki, Finland

In Finland, municipalities are the most important organisers of social and health care services. The decline of public economy has recently limited possibilities for municipalities to account for the services. There is an increasing interest to find out ways to improve the efficiency of welfare service production. The special challenge in the social welfare field is the lack of practical tools to evaluate efficiency, i.e. the costs, outcomes and effectiveness. There is an urgent need for the evaluation tools, and creating the tools is extremely challenging.

The objective of our project ‘Efficient and effective welfare service production’, is to develop practical evaluation tools for that purpose. The tools are required to improve efficient production of welfare services in municipalities.

The development of the evaluation tools is based on the scientific method of economic evaluation. The method is used in many sectors. There are sophisticated applications e.g. in health economic and especially in pharmaco economic evaluation. The scientific method raises certain requirements, which demand negotiation in practical municipal context. An example of these challenges is e.g. the demand of the theory to take into account of all costs for all stakeholders. In contrast, the municipality's interest lies in the costs for themselves. Another example is the target to measure effectiveness, which is difficult to define and measure even theoretically in the social welfare field.

We are carrying out the developing project together with two municipalities, Pieksämäki and Mikkeli, in South Savo in Finland. The funding comes mostly from the European Social Fund. We started to develop the evaluating tools to pilot services, home care for the elderly and sheltered housing, selected by the municipalities. Comparable costs will be defined, as well as relevant and measurable outcomes indicating efficiency. A simple and practical way to collect the data will be constructed.

Keywords: Social welfare field; Economic evaluation; Evaluation tool

Evaluation and stakeholding development

M. Reynolds
Open University, Communication and Systems, MiltonKeynes, United Kingdom

How is “value” or “worth” decided and what are the consequences of that for the design, data collection, analysis and reporting of evaluations? Critical Systems Heuristics (CSH) is widely known amongst systems practitioners and is deeply evaluative. Arguably it is one of the most powerful evaluation frameworks yet developed, yet ironically it is almost unknown in the mainstream evaluation field. Deceptively simple (12 questions asked from two orientations) the complexity of what it does reveals itself layer by layer. Critical systems heuristics (CSH) is a toolbox of 12 basic questions embodying a set of principles for evaluating the (i) built-in values, (ii) power structure and (iii) knowledge-base, associated with a specified focus (or ‘system’) of interest. At the same time CSH evaluates (iv) the moral basis on which a system operates as considered from the perspective of both beneficiaries and victims of the system. Whilst scientific data, including statistics, might be usefully channelled to support the output of a CSH evaluation, the overall evaluation is a qualitative exercise. CSH is nested in a constructivist tradition of systems practice primarily aimed towards collaborative improvement of, and developing responsibility over, the situation in which a system is embedded. Drawing on a significantly updated version of CSH co-authored by Reynolds and the originator of CSH, Professor Werner Ulrich, this paper describes how CSH can foster a concise mapping of stakeholder involved with any evaluation including beneficiaries, decision makers, experts and ‘victims’ (those affected by but not involved with a system of interest). It maps out what’s at stake in relation to each stakeholder group and the type of stakeholding anticipated for each group. The author argues that a prime feature of CSH is that the evaluation process promotes possibilities of stakeholding development for each stakeholder group as against stakeholding entrenchment.
How useful are user satisfaction surveys? A patched-up design for impact evaluation

L. Lindgren
School of Public Administration, Gothenburg

How satisfied are you with the health care services received? How likely is it that you choose the same service provider next time you need health care? Over the last decade patient (or consumer/user/client) satisfaction have gained widespread recognition as a measure of quality in many public sector services, including health care. Patients are asked to give feedback through surveys on different dimensions of services offered. Results are presented in databases such as Patient UK, US MedlinePlus, Patient UK, Open comparisons in Sweden, or the international Health Consumer Powerhouse. Patient satisfaction survey results are also commonly incorporated in structured systems designed to enable comparisons and quality ranking.

Although there is i) little evidence that survey results are effectively used by managers to improve the quality of health care, and ii) much evidence that comparative measurement can be detrimental to quality, governments worldwide continue to invest considerable resources and political confidence in the implementation of recurrent patient satisfaction surveys. In addition, a substantial industry of official regulators, academic institutions and commercial companies have developed who use patient satisfaction surveys and the comparative performance systems into which the surveys are incorporated partly to pursue their own ends.

Against this problematic background, and inspired by the drive for and increased attention paid to the issue of credible evidence in evaluation in recent years, this paper presents and discusses a quasi-experimental research design for evaluation of the utilization and effectiveness of patient satisfaction surveys. The study (which will commence in August 2010) focuses on a sample of primary health care centers in the region of Vastra Gotaland in Sweden. The main evaluation question is: Have the results of the patient satisfaction survey of 2009 shown any improvements on the dimensions where the primary health care centers scored poorly in 2007, and if so why?

Keywords: Health quality; Patient satisfaction surveys; Quasi-experimental design
Ethical Issues in evaluation practice

The African Evaluation Guidelines: a Cultural Dimension
M. Mbaireweaye, S. Jacob
ULAVAL, Political science, Quebec City, Canada

Most publications in evaluation focus on technique and methodology (Weiss, 1987; Chelemsky, 1987). In doing so, one omits to recognize that evaluation is a moral critic that involves values (Schwandt, 1991). This explicit recognition of values turns towards a cultural orientation of evaluation. The development of evaluation standards, since 1980, reinforced this new orientation. The most well celebrated among these standards and the one that caught our attention is the Joint Committee's Evaluation Program Standards. It consists of program evaluation standards (PES) carried out by the African Evaluation Association and named as the African Evaluation Guidelines (AEG). The transfer from individualist society to collectivist society has been criticized by Taut (2000), as impossible or considered being difficult by Russon (2004). As African societies are presumed collectivists (Hofstede, 2001), and because African Evaluation Association has adapted the Joint Committee’s Evaluation Program Standards within African context (Rouge, 2004), our statement aims at knowing whether Joint Committee’s Evaluation Program Standards adopted by the African Evaluation Association followed African’s collectivists cultural principles.

Our communication was based partly on the report of the debates occurred during the development of African Evaluation Guidelines (AEG) and on the other hand based on telephone surveys done on 8 (eight) key members of the African evaluator’s team and one foreigner non African expert that had followed the process.

Results show that transfer of the Joint Committee's Evaluation Program Standards within an African context by the African Evaluation Association was raised from a dynamic superimposition dynamics rather than from a replacement dynamics. Therefore, the cultural orientation underlying African Evaluation Guidelines is a hybrid one both individualistic and collectivistic in nature.

Keywords: Ethic; Values; Evaluation standards; Culture in evaluation; Quality

Inter-cultural evaluation. The challenge of the new international projects and informational space.
L. Korporowicz, S. Jaskula
Jagiellonian University, The Faculty of International and Political Studies, Krakow, Poland
Pedagogium, The Chair of Informational Culture, Warsaw, Poland

Background
There is a great progress in expansion both small and great international projects that demand special analysis of the social and cultural context of their realization. Enlargement of the range and functions of these projects creates qualitatively new situation in the field of their quality assurance, social rationality, adequacy on the national and local level.

Objectives
In order to know the value of these projects it is necessary to adjust evaluation methodology and evaluation criteria to local and specific demands as well as disclose processes of their inter-action, changes and mutual influence. On the other hand such evaluation should be open for new possibilities in expansion of the social and informational space. The purpose of the presentation is to propose interactive and sensitive to cultural environment, dialogical and responsive type of evaluation From this point of view intercultural evaluation is a kind of cultural learning, discovering and a chance for communication.

Results
As a result of the presentation we would like to describe an explanatory model for evaluative cultural misunderstandings but first of all for designing evaluation as a cultural discourse with respect to core values and identity of the particular groups, to their ability to cultural change and development.

Application
The concept of the inter-cultural evaluation should be applied in all international and cross-environmental projects of evaluation, especially in qualitatively oriented studies in multicultural communities or organizations. There is a great need for evaluational inculturation in order to reduce both mental and social resistance as well as stimulate evaluation activity to more interactive strategies and practice.

Keywords: inter-cultural evaluation; Evaluational inculturation; International projects; Informational space; Cross-national studies
Ethical questions in the process of development aid and cooperation

K. Princova
Palacky University, Faculty of Theology, Olomouc, Czech Republic

There are many ethical questions arising during the whole process of delivering of international aid and realisation of international development cooperation. The most important ethical questions relate to the local community, local people. There originates the most serious risk, that the whole aid is sometimes not effective. The reasons of it could be very different. For example the recipients do not understand the purpose of the development projects properly, so the realisation does not have the impact which was intended. A good part of the aid does not reach the recipient community at all.

The main problem, the international aid strives to reach is hunger and poverty. The development aid nevertheless very often does not eliminate poverty but in many cases the international aid and cooperation could even enhance the poverty of different groups and communities and increase the differences between them.

Is it responsible to plan, promise and negotiate without being convinced of the value of what are we doing? Could it be that the evaluation of the project is O.K., but in spite of that the aim was not reached? Is it possible to change this situation? Do we have any instrument, which could reach the recipient more properly?

The aim of this lecture is to proof that non ethical behaviour during the whole process of international aid cannot effectively reach the goal of it.

Keywords: Ethics; Responsibility; Effective
Governance of RDTI policies in a complex environment: Agencification and the interaction of federal ministries and funding agencies in Austria.

S. Ruhland1, I. Fischl1, S. Mayer2

1 Austrian Institute for SME Research, Evaluation InnoTech, Vienna, Austria
2 Austrian Institute for SME Research, Evaluation, Vienna, Austria

In 2008/9, the Austrian Institute for SME Research – as part of a consortium – conducted the evaluation of the Austrian system of public R&D funding and support. One of the main evaluation questions was about whether or not the interaction between the strategic (ministries) and the operational (funding agencies) level of national RDTI policies performed well and as intended.

In general, the division of labour between ministries and agencies is more or less clear (on paper) and is perceived as being comprehensible, too. However, in the day-to-day business of stakeholders in Austrian RDTI policies it is much more blurry and individuals – depending on their respective importance, access to budgets etc. – tend to extend and reinterpret their office. Furthermore, the accumulation of both operational and strategic knowledge in the funding agencies leads to an unbalanced distribution of information in the system of public R&D support. The broadening of responsibilities therefore, additionally applies for organisations such as funding agencies as whole. Ultimately, the system is affected by individuals with long careers working their way up or ‘crossing the divide’ without ever obtaining an external perception.

Although the system proves to work fine, its performance is achieved despite the way it is structured and operating. A lack of strategic guidance and trust was observed, which lead to recommending to introduce a process of ‘earned autonomy’ that guides a review of the division of labour and its everyday implementation, organises the return to everyone’s primary tasks by means of a step-by-step process shifting tasks including revisions of the performance and the process itself as well as the possibilities to revoke tasks granted in case of bad performance. Eventually, this process supports the development of trust that is needed for a better cooperation in the future for the sake of the public Austrian RDTI support system.

Keywords: Governance; RDTI policy; Agencification
Evaluation of Public Research and Development Policy for New Drug Development using System Dynamics Approach

S. J. Yoo1, B. Gill1, G. Park2, J. Choi3

1 Korea Institute of Science & Technology Evaluation and Planning, Division of R&D Program Evaluation, Seoul, South-Korea
2 Korea Institute of Science & Technology Evaluation and Planning, Office of Science & Technology Policy and Planning, Seoul, South-Korea
3 Ministry of Strategy and Finance, Performance Management Division, Gwacheon, South-Korea

New drug development is a major area for commercialization of biotechnology research and development (R&D), which requires longer time and huge financial investment. And that the success rate of new drug development is getting difficult, even though most of multinational pharmaceutical companies put a lot of resources to replace their current blockbusters. Back in Korea, most of pharmaceutical companies run short of resources to tackle new drug development issues. Therefore, government invests public finance to support poor R&D environment of pharmaceutical companies.

In the present study, we try to find the optimum investment strategy and allocation of public finance into R&D for new drug development, since we noticed that inefficient investment has been made to support preclinical and clinical trial of new drug candidates of pharmaceutical companies. Therefore, we tried to delve into R&D policy and investment strategy for new drug development of the government. Here we applied system dynamics model to evaluate the past public policy and investment strategy for new drug development and subsequently to suggest the optimized budget allocation at each development stages. In this model, we collected practical data of average budget and success rate for each phase and inventory size of the previous phase, etc. We modified the previously suggested model from our colleagues to adjust to the current situation and analyzed the data using Vensim® simulation software.

We found that budget would be better to allocate later phase, for example, clinical phase instead of preclinical phase, or phase 2 of clinical trial instead of phase 1, etc. after simulating with practical data obtained from the pharmaceutical companies.

Taken together, we’d like to suggest that government establish a budget allocation strategy according to needs of the companies and facts they have such as inventory, success rate and average budget of each phase of development.

In-depth evaluation of Research and Development program for various energy sources substituting for fossil-fuels

S. Nah, J. Hong

Korea Institute of Science & Technology Evaluation & Planning, Division of Evaluation & Coordination, Seoul, South-Korea

As global warming becomes threat to human life, the world is now suffering from and facing unprecedented ever-increasing crises with respect to energy consumption and the depletion of natural resources as well as economic hardship. Korea is no exception and ranked the 1st among OECD member countries in terms of its rate of increase in greenhouse gas emissions and, as one of the 10 largest energy consumers in the world, Korea relies on overseas imports for 97% of its total energy consumption. To overcome these difficulties, our government has implemented many R&D projects to substitute new and renewable energies for fossil-fuels, under governmental policy of “Green Growth”. Therefore, lots of public finances are invested by government to support the projects, which also means need to evaluate outcomes of the programs.

In the present evaluation, we make an attempt to find an optimized investment strategy, allocation of financial resource, and regulation of programs, since we noticed that it is difficult to decide certainty of success, especially in the green technology R&D whose boundaries are various. Thus, we tried to investigate R&D policy and investment strategy under such angles as relevance, efficiency, effectiveness, and overall program system. In the analysis, we used top-down approach and MARKAL as a model of energy optimization.

In conclusion, we found that Carbon-dioxide (CO₂) reduction should be emphasized and accordingly we’d like to suggest that governmental investment strategy should be established, considering the level of reduction, technological competitiveness, supply possibility, and overseas marketability, together with national energy security.
Trends and issues in European university programs in evaluation

A panel proposed by the USPE-collaboration

On this panel, the most pressing trends and issues concerning European university programs in evaluation will be discussed. Ahead of the conference, the surveys of 2004/2005 and 2008 (cf. Beywl & Hanich, 2007; 2008) will be replicated in order to obtain an updated overview of the characteristics of the programs (e.g. entry requirements, duration, attendant type, price, core subjects, learning outcomes, assessment types, graduates). The results of the survey will briefly be presented and handed out to participants of the session. Then the representatives of different study programs will speak about current challenges and future trends concerning their programs (e.g. undergraduate vs postgraduate education, degree of field specific training, didactics of evaluation, certification). To conclude, there will be debate among panelists and open discussion involving the audience to share ideas about how to make university programs in evaluation work, and what the challenges are for the future of such programs.

Proposed speakers on the panel:
- Maria Bustelo (University of Madrid, Master in Evaluation of Programmes and Public Policies)
- Peter Dahler-Larsen (University of Southern Denmark, Master in Evaluation)
- Verena Friedrich (University of Bern, Diploma & Master of Advanced Studies in Evaluation)
- Stefanie Kihm (Saarland University, Master Studiengang Evaluation)
- Heike Steckhan (University of Bonn, Master of Evaluation)
- Jacques Toulemonde (Eureval and Lyon University, Master Monitoring and Evaluation of Policies and Programmes)

Chair:
The session will be prepared and chaired by Peter Dahler Larsen (University of Southern Denmark) and Verena Friedrich (University of Bern), on behalf of the European Collaboration on University Study Programmes in Evaluation (USPE), a network of six European university programmes delivering master degrees in evaluation.

Keywords: Training in evaluation; Study programs
Evaluating EU Educational policies & systems

In search for participation and transparency- Evaluating organization for reform of national course syllabi in compulsory school

L. Niklasson, C. Carlhed
Mälardalen University, School of Education Culture and Communication, Eskilstuna, Sweden

The aim with the article is to present and discuss how a evaluation was created and carried out during 2009 and 2010 concerning organizing a process to change course syllabi for compulsory school in Sweden. The background is that international and national evaluations have shown that results from Swedish schools are not improving. The pupils are not reaching the goals and the teachers criticize, among other things, the course syllabi. As a response to these results a national reform regarding curricula and course syllabi for compulsory schools was initiated from the government, the reform was called School 2011. The National Agency for Education was aware of the forthcoming reform and when the directives came from the Ministry of Education and Research during Spring 2009, an organization was already designed for the development of the new course syllabi. During Autumn 2009 Mälardalen University got the opportunity to design an evaluating that was carried out late 2009 and Spring 2010. The design for the evaluation had to consider the vision the National Agency for Education had with the organization which was to increase transparency and participation during the process, compared with earlier revisions of curricula and syllabi, besides creating new course syllabi as a product. The data collection during the evaluation comprised text analysis, observations, interviews and questionnaires. The respondents were those who worked with the first version of the course syllabi, workgroups, reference groups, researchers, teachers and the staff at the National Agency for Education who wrote the last version. A conclusion that can be made from the design of the evaluation and the data collection is that transparency and participation can be discussed both from a political and a practical perspective.

Keywords: Reform; Compulsory school; Transparency; Participation; Evaluation

A boost for teachers – evaluating a national programme for professional development

L. Westin¹, R. Hort², A. Lundberg³
¹Swedish Agency for Public Management, Stockholm, Sweden
²Swedish Agency for Public Management, 3, Stockholm, Sweden
³Swedish Agency for Public Management, 1, Stockholm, Sweden

Background
National and international studies show a decline in children’s achievements in Swedish schools. The professional qualifications of teachers are regarded as a crucial factor in addressing this problem. Therefore, in 2007, the Swedish government launched a national programme – The Boost-for-Teachers Initiative – to upgrade teacher qualifications. The programme includes government subsidies to enable teachers to take part in further education programmes at university level, while receiving 80 per cent of their current salary. The aim is that a quarter of all teachers will participate in the programme by the end of 2011, when the programme will be phased out.

Evaluation design
The Swedish Agency for Public Management has performed an evaluation of this programme, on behalf of the government. The evaluation deals with many aspects of the programme. Central questions are: In what ways have teachers upgraded their qualifications? How do teachers experience the content and the quality of the further education programmes provided? How well has the National Agency for Education fulfilled the task of implementing the programme? Part of the assignment has also been to design an evaluation of the programme on a long-term basis. Several methods such as interviews, focus groups and questionnaires have been used as the basis for the evaluation.

Challenges
Some challenges met in this evaluation are of general interest and therefore deserve a discussion. For example, several changes have been made in the programme since it was launched. How does one evaluate a moving target? Another challenge is to evaluate a programme with different, and sometimes conflicting, objectives. A central issue is what assessment criteria should be used to evaluate the results of the programme. Also, is it meaningful to evaluate a programme from a national perspective given that the educational system is the responsibility of municipalities?

Keywords: Education; Conflicting objectives; Moving target
Development of the procedures for evaluating the education policies inside the national educational systems

M. Roman
National School of Political and Administrative Studies

Education and the professional formation occupy a central position within the European Union’s agenda for the year 2020. The financial crisis proved once more the necessity of obtaining and developing the professional abilities necessary for creating new jobs and prosperous societies. Thus, EU member states have the responsibility of reforming their own educational systems and adopting new public policies inside the sphere of education, having as main goal the complex development of their citizens – personal development, as active citizens, and as capable professionals – in order to best respond to the demands of their societies.

The process through which the public policies inside the educational system are analyzed, thus obtaining a feed-back for the results, impact and expectations for them, is their evaluation process. Evaluation is extremely important when talking about professional training and a performing education. Most EU member states are currently practicing centralized evaluations for obtaining key abilities. Most current evaluation methods focus on memorizing and obtaining information, disregarding the abilities and the capacity to use the acquired abilities. Considering this particular context, it is mandatory to use complex evaluation methodologies, and one needs to benefit from the experience of the countries using such procedures.

What are the questions to be answered by the evaluation of the aims set through adopting new education public policies? What are the evaluation criteria that needs to be considered? What are the fundamentals when talking about an efficient evaluation? How are the education public policies evaluated in EU member states? What can the member states learn from one another in the process of developing the evaluation procedures for the education policies in the national educational systems? This paper will provide answers and proposals regarding the methods for evaluating the education public policies inside the general national educational system.

Professional and social success are directly connected with the methods for the evaluation of education and professional training.

Keywords: Evaluation, Society, Education policies
Evaluation and Policies: Implementation and effects

The policy for monitoring and evaluation of primary health care in Brazil, 2003–2006: contextualizing its implementation and effects

E. Felisberto1, E. Freese1, C.K. Almeida Alves1, L. C. Albuquerque Bezerra1, I. Samico1, Z. Hartz4
1 Centro de Pesquisas Aggeu Magalhães/Fundação Oswaldo Cruz – CPqAM/Fiocruz, Public Health Department, RECIFE-PE, Brazil
2 Instituto de Medicina Integral Prof. Fernando Figueira – IMIP, Research Department – Research Group of Health Evaluation, RECIFE-PE, Brazil
3 Instituto de Higiene e Medicina Tropical – Universidade Nova de Lisboa, Board of Directors, Lisboa, Portugal
4 Instituto de Medicina Integral Prof. Fernando Figueira – IMIP, Research Department, Lisboa, Portugal

This research presents a study of the implementation of the Policy for Monitoring and Evaluation of Primary Health Care in Brazil during the period of 2003 to 2006. It estimates the implementation degree of the institutional mechanisms and the intervention, and evaluates its influence in the effects produced from its interaction with political and organizational context. It was carried out an evaluative research (Implementation Analysis). It is a case study with interposed levels of analysis and triangulation of methods. To estimate the implementation degree, it was considered the institutional mechanisms produced that are representative of the components of the intervention logic model. It was carried out 20 interviews and 29 documents were analyzed. The intervention was classified as implemented in an advanced degree, with 78% of actions being done. Institutional political context was characterized by a conjunction of political institutional decision, financing resources, means of technical and organizational strategies of human resources qualification for evaluation towards the institutional capacity building. It was found a significant amount of no-intentional effects as a consequence of various mechanisms used to the implementation of the intervention. The Policy discloses a great effort towards actions integration on the institutional level and to the decentralization of evaluation actions. In this way, two characteristics are relevant: the development of technical capacity by formative evaluative processes and the focus on the empowerment of the Health State Secretariats.

Keywords: Implementation analysis; Evaluation; Health program evaluation; Primary health care

Improving legislation and policies – evaluation as a tool for Smart Regulation in the EU

S. Hector, S. Jakobsen
Directorate C Secretariat General, Evaluation Unit, Brussels, Belgium

The European Commission uses evaluation for evidence-based decision making to plan, design and implement EU policies. For more than a decade it has been widely used for the expenditure instruments of the EU. The new Commission established in 2010 has made evaluation an important part of its working methods, taking forward the achievements on Better Regulation of the previous Commission. Ex post evaluation along with impact assessment, simplification and reduction of administrative burden will be key tools to ensure that EU law provides a framework for a fair and effective single market, which at the same time protects workers, consumers, health and the environment.

The Commission is therefore stepping up its efforts on evaluation, ensuring that also regulatory measures will be subject to systematic evaluation. Evaluation of the effectiveness of existing policies is a fundamental step in developing new initiatives. A real understanding of how far existing policies are producing their intended results must allow us to correct shortcomings and to design new initiatives to have a real impact. All significant proposals for a revision or new measure should be based on an evaluation of the policy framework already in place. This will be the basis for proposing evidence-informed adaptations, which are analysed in the ensuing impact assessment.

To this end the Commission is planning a range of evaluations for the years ahead, closely linked to its work programme of initiatives to be presented to the European Parliament and Council. These regular evaluations will be complemented in policy sectors with a complex and fragmented regulatory framework; there the Commission intends to launch pilot evaluations looking at an entire policy sector to identify overlaps, gaps, inconsistencies, obsolete measures, excessive administrative burdens, etc.

Keywords: Smart Regulation; Better Regulation; Regulatory Evaluation; EU legislation; Fitness check


F. L. Leeuw1, Q. Zhai2
1 Maastricht University, Faculty of Law, Maastricht, Netherlands
2 National Center for Science and Technology Evaluation, Beijing, China

Hong Houqi and Ray C. Rist have described the situation in China regarding evaluations before 2000 in their article on Evaluation Capacity in the People’s Republic of China: Trends and prospects (published in the: The international Atlas of evaluation, Transaction Publishers, 2002). Over the last decade great changes have taken place in the field of evaluations in China, especially in the public sector. The paper will use a literature review and a comparative analysis to introduce and explain the character of the changes, such as the evaluation policies of the different layers of government in China, the evaluation research and the evaluation practice, both at the central and provincial level. The paper will also pay attention to the question how these changes can be explained. Finally, the paper will provide the challenges for evaluation and evaluation policy in China.

No conflict of interest

Keywords: Evaluation culture; China
Strengthening evaluation thinking through external evaluation: Rewards and risk

C. Duggan¹, F. Carden¹, I. Davies²

¹International Development Research Centre, Evaluation Unit, Ottawa Ontario, Canada
²European Evaluation Society, Paris, France

Can the goals of accountability, learning and legitimacy be combined in program level evaluation in publicly-funded international development organizations? Evaluation in the “public interest” suggests that, in order to be effective, evaluation should accommodate the interests of multiple stakeholders: Commissioners, evaluators and stakeholders. Balancing these different sets of interests can generate tensions between participation, timeliness and use of findings.

The International Development Research Centre (Canada) is currently piloting a new method for conducting program evaluations to provide advice to the Centre’s Board of Governors on progress achieved and challenges that remain. In combining self-assessment with external expert opinion, IDRC intends to meet accountability goals while enhancing learning and strengthening a culture of evaluative thinking among Centre staff. The new method is currently being piloted with a number of programs, including its Evaluation Unit. This panel will consist of three paper presentations: One paper will discuss the method from the perspective of the subject of the evaluation; another paper will examine the evolving experience from the perspective of the evaluation manager; and the third paper will offer insights on the promise and pitfalls of the method from the perspective of the evaluator. While it is too early to offer any conclusive comments around the success or failure of this experiment, the panelists will share observations on the evolution of the process, strengths and weaknesses of the design with relation to the different goals, and initial feedback received from key actors in the evaluation process (subjects and users of the evaluation).

Keywords: Self-assessment; External review; Evaluative thinking
Evaluation and capacity development

Advocacy evaluation – is it really that different?

E. Pritchard
World Vision Australia, Program effectiveness, Melbourne, Australia

In recent years there has been increasing interest and involvement by evaluation practitioners in the field of advocacy. This growth has paralleled and in part resulted from a maturing of understanding in the international development sector about the critical role of advocacy for achieving sustainable transformation in communities.

Advocacy work is characterised by some unique features which necessitate additional thought and adaptation of the evaluation practices usually engaged by non-government organisations. Compared with other areas of international development, the environment of advocacy work is characterised by more uncertainty, less control and changes (both anticipated and unanticipated) of both great rapidity and tremendous slowness. These qualities require project strategies and designs of exceptional sensitivity and responsiveness. Thus, continuous self-evaluation is critical to the success of advocacy work, arguably more so than for other international development activities.

Based on recent evaluations of World Vision advocacy projects promoting human rights in different countries and in restricted political contexts, this presentation reflects upon questions that arose:

• What are the benefits of reviewing the theory of change?
• How much effort should be given to verifying claims of success or wins?
• How useful are key informant interviews in restricted political contexts?
• Setting indicators – how smart is SMART?
• Does advocacy evaluation really involve more ‘intangibles’ or is it an epistemological question of what constitutes evidence?
• What is the role of an external evaluator where advocates have strong evaluation capacity, are up to date with changes in context and the implications of these for their planned activities and approach and know how to powerfully share their story?

This presentation aims to promote discussion about how evaluations can be designed to best support advocates working with and within the complexities of developing countries.

Keywords: Advocacy; International development; Evaluation design; Restricted political context


R. Suelzer
GTZ retired senior advisor evaluation, independent, Berlin, Germany

Background

Capacity Development (CD) is of highest significance in development cooperation (DC) as the ultimate rationale justifying the means. In Europe the foundation of the European Centre for Development Policy Management (ECDPM) in 1986 focussed since on fostering CD initiatives and in 1997 the Canadian CIDA went into further operational details with “The Design and Use of Capacity Development Indicators”.

Since then, related publications and websites have almost obscured the focal concept of CD: managing one’s own affairs successfully – without continuous external support in the long-term. The demanding concept of rigorous evaluation and impact evaluation (IE) and the focus on rather short-term results based management (RbM) leads to a neglect in search of capacities sustained long after phasing out of foreign-aided programmes.

Objectives

The study attempts to identify the factors that may lead to large scale social and economic development programmes and projects which finally are not resulting in long-term sustained social change thus not effecting in the long run effective change in the regular operations of concerned organizations (administrations).

Methods

Retroactive meta-evaluation of outstanding ex-post evaluations is used and ex-post evaluations are analysed which have a sound evidence-based data set and a sufficient distance to the concerned development programme.

Results

The meta-evaluation of an outstanding, comprehensive Ex-Post Evaluation of BMZ (Germany) in 11 volumes of the Area Development Programme West Pasaman, Indonesia, and other evaluations assessed lead to the conclusion that the Indonesian and the German side, both, focussed on short-term, visible “impact” but rather not on “sustained institutional change” or “sustained capacity”.

Conclusions

The concept of Capacity Development – and its subsequent and ongoing impact evaluation – requires a design and implementation different from the short-term impact oriented development measures so common. And it requires a different approach in the design of evaluations, both, formative and summative.

Keywords: Retroactive meta-evaluation; Capacity development in evaluation; Methods of ex-post evaluation; Lack of CD concepts in evaluation design; Lack of CD concepts in programme design
Consequences of Evaluation Capacity Building: A multiple case study

J. B. Cousins¹, C. Elliott¹, L. Morra Imas¹, B. Levine¹, R. Rist¹

¹University of Ottawa, Faculty of Education, Ottawa Ontario, Canada

Research on the consequences of ECB initiatives has considerable potential to help inform our understanding of nature of impact as well as ECB processes and strategies that are likely to lead to sustainable capacity development. Yet published research on evaluation capacity is relatively limited in quantity has mostly taken the form of experience-based narratives, written by evaluators working on evaluation capacity building (ECB) initiatives (Cousins, Goh, Clark & Lee, 2004).

In this paper we report the results of a multiple case study of the consequences of development evaluation training and capacity building. Sites were selected by virtue of sustained experience and relationship with the International Program for Development Evaluation Training (IPDET). IPDET is co-sponsored by the World Bank and Carleton University and has been offered annually in Ottawa, Canada since 2000. The multiple case study is part of an impact evaluation of IPDET and is conceptually guided by research on evaluation capacity (Cousins et al., 2004). The questions guiding the research are: (1) What are the consequences of evaluation training in terms of knowledge development, the capacity to do and use evaluation, the integration of evaluation into organizational culture, and networking within the evaluation community, broadly defined? (2) What aspects of training are salient in explaining patterns of outcomes? To what extent are training outcomes mediated by local context?

Five sites were selected for study. Three of the sites are country-level sites where a substantial number of participants have attended IPDET training in Ottawa over the years. Two sites are organizational-level sites having evaluation capacity building as part of their mandate. Data collection included interviews and document analysis. The cross-case findings will be discussed in terms of their implications for understanding the consequences of ECB and considerations for the design and delivery of development evaluation capacity building initiatives.

Keywords: Capacity building; Training; Impact; Multiple case study
Evaluation use, evaluation impact II

Fostering evaluation use: lessons from a large systematic review on knowledge transfer

D. Contandriopoulos¹, A. Brousselle²
¹ University of Montreal, Faculty of Nursing, Montreal, Canada
² University of Sherbrooke, Faculty of Medicine, Longueuil, Canada

Evaluation can serve a purpose only to the extent that its results are used. Use is thus both a core objective of evaluation practice and an essential dimension of all evaluation theories, according to the framework proposed by Shadish & al. (1991). The relative importance of use as compared with the four other core dimensions (social programming, valuing, knowledge and practice) of evaluation theories and the influence it has on the evaluation process will, however, differ from one evaluation approach to another.

From a systematic review of the field of knowledge exchange and utilization and in-depth analysis of 205 papers on the subject, we identified two core dimensions that structure knowledge use: the degree of issue polarization and the balance of cost-sharing between producers and intended users. Using this typology, we contrasted the results of the systematic review with the position on utilization of some well-known evaluation theories.

At a first level, our results show that evaluation theories actually differ sharply on the expected polarization of the issue and the expected distribution of cost sharing between producers and intended users. At a second level, we suggest that the fit between the actual context of use and the expected conditions put forward by evaluation theories is a credible predictor of evaluation use. This leads us to three important conclusions: 1) No evaluation theory, strategy or practice intrinsically fosters use. 2) What is crucial is the fit between the chosen approach and the actual characteristics of the context. Finally, 3) some evaluation approaches will limit themselves to high use contexts, but this will strongly structure the issues, methods and results of the evaluations done.

Keywords: Results use; Evaluation theory; Knowledge exchange

Strategies for Learning from Evaluation in a Post-heroic World

M. Vahlhaus
German Technical Cooperation (GTZ), Evaluation, Eschborn, Germany

Forget people – it’s the organisations that matter: Learning literature has been obsessed with the way individuals learn for most of the industrial era. There is little or no hard evidence on how organisations learn. Organisations learn by fixing learning in new or adapted structures, processes and rituals.

Learning from Evaluation requires learning at individual, organisational and network levels to be effective and sustainable. This is to say: Post-heroic (i.e. post-individualistic) learning models must develop convincing models of how learning from evaluations happens on individual and organisational levels and especially how these levels interact and influence each other. Organisations learn different from individuals: Organisations are basically unwilling to learn. They are built to maintain certain routines (that is the justification for their existence). Any strategy that requires organisations to change brings the organisation into a fundamental conflict of objectives. Organisations learn by incorporating knowledge into their structures, processes, rules and rituals and by making them intelligent.

Learning from evaluation has to take these aspects into consideration if we want evaluations to be effective and to have an impact. GTZ (German Technical Cooperation) has implemented an independent evaluation unit in 2006. Since then the evaluation unit together with other units such as planning and development as well as knowledge-management has made several efforts to meet these challenges in the further development of respective evaluation instruments and processes. E.g. by conceptualising decentralised evaluations (purpose/standards/quality assurance), by focusing independent evaluations and their analysis on sectors, by developing and organising new learning formats such as learn cafés.

Keywords: Learning; Organisational Development
Constraints on the Use of Evaluation in Policy Making: Lessons from Africa

A. E. Gakusi1, A. N. Sindzingre2

1 African Development Bank, Operations Evaluation Department, Tunis, Tunisia
2 National Center for Scientific Research (CNRS) – University Paris 10 – School of Oriental and African Studies (SOAS) University of London, Department of economics, Paris, France

The paper aims to analyze the constraints that prevent the use of evaluation, with a focus on African countries. It is commonly agreed that the role of evaluation should be to generate high-quality and highly relevant evaluation findings for use in program decision making and improvement. However, these two steps in evaluation are in most cases beset by daunting obstacles that lessen the relevance, credibility and usefulness of evaluation findings, and hence, impede their efficient implementation.

This paper contends that the development evaluation process faces particular constraints that are inherent in the political economy of aid agencies and that of recipient countries. Both entities may pursue dissimilar objectives because of diverging preferences and interests. This may create perverse incentives, including moral hazard. The use of evaluation is also hampered by a set of factors, in particular: (a) lack of proper incentives both in donor agencies and developing countries; (b) weak research methods and inadequate information; and (c) different organizational rules and value systems. Furthermore, the aid system is characterized by weak accountability both on the donors’ and recipients’ sides. On the donors’ side, in contrast to the national budget where the government is accountable to tax payers, in particular through elections, aid agencies create a disconnection between tax payers and beneficiaries. On the side of beneficiary countries, accountability is also weak for different reasons, notably the nature of the political system, the low level of education of the population, and poor institutions. The lack of use or underutilization of evaluations reflects the intrinsic tensions that exist at the different levels of the aid process. This paper will illustrate these tensions in the development evaluation system by drawing on examples from the African context. It will also provide reflections on ways that would enable an exit from these tensions.

Keywords: Evaluation use; Development; Sub-Saharan Africa; Accountability; Aid
Methodology and sectoral policy

Evaluation research as instrument to improve child and family policies

G. Moro, C. Balenzano
University of Bari, Department of Psychology, Bari, Italy

Background
Adozione mite is a particular form of open adoption implemented by Juvenile Court of Bari (South-Italy). Differently by traditional adoption, it isn’t characterized by a drastic interruption of the bonding with biological families. Who strongly believe in adozione mite argue that this instrument can represent a convincing response to permanent semi-abandonment social emergency.

Objectives
By means of analysis of families paths and outcomes, this research intends to evaluate if this instrument can ensure families and children wellbeing.

Methods
The first exploratory study used the method of archival research to describe characteristics and paths of 119 children and families. Moreover, we have implemented a multi-stakeholder evaluation study to identify evaluation criteria of adozione mite, analyzing by focus group discussions stakeholders’ representations (Moro, Cassibba, Costantini, 2007). Finally, we have evaluated the effectiveness of adozione mite thought an outcome evaluation study on adoptive parents and sons.

Results
Data analysis shown that in 12.6% of cases, children have psychological syndromes and join foster family when they are school-age. Effectiveness criteria shared by stakeholders’ group are: children’s well-being, a positive relationship between birth and foster family and between children and birth family. The only failing criteria identified is the “psychological distress of the child”. Outcome evaluation study is still ongoing. Preliminary data analysis show that the frequency of contacts between adopted children and birth parents is generally low; the relationship between children and adoptive parents is generally very good (67.4%) while the relationship between birth and adoptive families is good or very good only in 25.6% of total cases.

Conclusion/Application to practice
When biological families are unable to provide to sons care and protection, adozione mite is applied to give to them the opportunity to live in families. Results from study 3 will be crucial to aim decision making process to support families and suggest innovative policies to face with children semi-abandonment emergency.

Keywords: Evaluation research; Child and family policies; New forms of adoption; Stakeholder evaluation; Focus group

A three step approach for evaluation of complex policy support programmes: methodological issues and lessons from the implementation

E. Caputo, C. Valmarana
Development Researchers’ Network, Evaluation, Roma, Italy

Background
The evaluation of complex policy support programmes is ever more a priority in different contexts (Development, Transition, EU). The focus is on two main issues:
• results assessment, and
• programmes and related policies’ contribution to results.

Objectives
To provide a conceptual framework for a close integration of process and impact evaluation. Process evaluation highlights the policy and institutional changes induced by a given policy support programme. Impact evaluation assesses the relevant social and economic changes occurred and their relation with the changes in the policy and institutional context.

The framework is applicable to the different contexts in which complex policy support programmes are to be evaluated, while specific techniques and methods for process and impact evaluation may change, according to the contexts and the availability of data.

Methods
The proposed methodology is based on a three step approach:
• the first step consists in an attempt to establish a linear relation of causality (contribution) between the inputs provided by the support programme and the changes in the institutional and policy environment.
• the second step consists in a regressional analysis aimed at establishing correlations between the targeted social and economic changes as actually occurred and the actual institutional and policy environment.
• the third step consists in the assessment of the transitive relation between the given support programme and the actual targeted results, such as: if (and how much) a (programme) has affected b (policies/institutions) and b has affected c (results), then a has affected c.

Results
• overcoming the limits of the theory-based evaluations, which overlook the results
• identifying the most suitable policy support packages to reach results.

Application to practice
The application by EC and DAC of this approach to large multi-donor country programmes in Tunisia and Mali will be illustrated and commented.

Keywords: Complex evaluation; Methodology; Impact evaluation; Policy evaluation; Case studies
Methodological challenges I

A practical approach to multi-sectoral outcomes and uncertainty in evaluation: active transport and voluntary travel behaviour change

Ian Ker
CATALYST (Consulting in Applied Transport, Access and Land use Systems)

Background
Transport policy is increasingly embracing areas beyond provision and management of infrastructure and services to include management of demand. Three key issues with this new paradigm are:

- Demonstrating the effectiveness of and outcomes from new ways;
- Valuing non-traditional outcomes of transport initiatives, such as health and social well-being; and
- Presenting multi-sectoral outcomes in ways that are appropriate for fragmented decision-making in the public sector.

A related problem can be a lack of assurance that such innovative initiatives can compare favourably with established programs using conventional evaluation methods.

Objectives, Methods and Outcomes
The paper describes an evaluation framework developed by the author for use as a decision-support tool and presents the outcomes of a range of applications.

The framework is designed for use in a range of evaluation contexts, and the paper will discuss methodological issues inherent in applying such a framework. This framework acknowledges lack of certainty of outcomes, especially as demonstrated through variability between initiatives, and provides graphical as well as numeric results.

Many of the newly-identified important outcomes, particularly those outside the usual areas of concern, are unable to be given robust quantitative values. However, where decision-making is a well-established process, it is important to bring the qualitative into that process, at least in the short term, rather than trying to establish new processes, which would be a long process in itself.

Assessing qualitative factors can usefully involve stakeholders rather than technical professionals.

Conclusion/Application to practice
It is essential to define the role of evaluation in public policy decision-making. The appropriate role is decision-support, not decision-making, and creating understanding rather than providing a spurious perception of accuracy.

This requires a flexible framework and ways of presenting results that are indicative and informative rather than spuriously precise.

Keywords: Multi-sectoral; Decision-support; Qualitative; Quantitative; Stakeholder

Evaluating patient and public involvement in health and social care research: methodological challenges

P. Carter, C. Jinks
Keele University, Primary Care Sciences, Stoke-on-Trent, United Kingdom

Background
Patient and public involvement (PPI) in health and social care research is strongly encouraged within UK and international policy. The policy is based on assumptions about the contribution that patients and the public can make to the research process and outcomes, besides participation being seen as key to good governance. Currently there are calls for evaluation of PPI but this is an emerging field with uncertainty around definitions and methodological frameworks.

Methodological challenges
First is the choice of paradigm. The rise of an evidence based policy and practice movement is premised upon a positivist methodology. This presumes outcome measures can be agreed and evaluation can determine “what works”. However, active participation in research may align with theories of participatory democracy and be viewed as a worthwhile end in itself. This leads to the second challenge, which is whether verifiable measures of PPI are appropriate? The value-laden nature of PPI means that normative (as opposed to “scientific”) judgements that appear in practice are also likely to emerge in evaluation. Third is the question of what the evaluating or object of study should be – what are the important observable processes and how should they be assessed? Fourth is the call for independent evaluation and whether this can be reconciled with a commitment to participation? A final challenge relates to the purpose of evaluation and who decides what is in the public interest. The policy of PPI in research has been implemented without evidence, so will evaluation make any difference?

Conclusion/Application to practice
Participation of patients and the public in health and social care research is now required. Evaluating how policy translates into practice offers a way forward for an evaluation design. This presentation sets out a series of challenges, arguing for a theoretical awareness of the politics of policy evaluation.

Keywords: Health; Research; Participation; Evaluation; Methodology
Research assessment in the Humanities: problems and challenges

P. Galimberti
University of Milan

Research assessment is going to play a new role in university governance moving from evaluation of results as a tool for resource allocation towards policy design.

In this respect “measuring” implies a different approach to quantitative aspects as well as to an estimation of qualitative criteria so difficult to define. Bibliometrics became so popular, in spite of its limits, just offering a simple solution to complex problems. The theory behind is not so robust but available results confirm this method as a reasonable trade off between costs and benefits.

Indeed there are some fields of science where quantitative indicators are very difficult to apply due to the lack of databases and data, in few words the credibility of existing information. Humanities and social sciences (HSS) need a coherent methodology to assess research outputs and the current projects are not very convincing.

The possibility of creating journal ranking at institutional, National or European level is not enough as it is raising the same bias as in the hard sciences and it doesn’t solve the problem of the various types of outputs and the different, much longer time of creation and dissemination.

The web (and web 2.0) represents a revolution in the communication of research results mainly in the HSS, and even evaluation has to take into account this change. Furthermore, the increase of open access initiatives (green and gold road) offers a large quantity of transparent, verifiable data structured according to international standard that allow comparability beyond national limits and is independent from commercial agents.

The pilot scheme carried out at the university of Milan for the Faculty of Humanities demonstrated that it is possible to build quantitative, on average more robust indicators, that could give a proxy of research outputs even in the HSS

Keywords: Research assessment; Humanities and social sciences; Bibliometric indicators
New tools for development evaluation

The usefulness of M&E Registry Book as a tool for enhancing Governance at Community level: Tanzania Experience

F. P. Madembwé
Prime minister’s office/Rural financial service programme, Mbeya, Tanzania

The importance of learning from Monitoring and Evaluation process has always been evident, but what is becoming increasingly clear is that those who are expected to use evaluation recommendations and lessons for improving performance should participate in the process of developing and implementing those recommendations. In line with efforts to rationalize the development of Decentralization by Devolution (D&D) in Tanzania – lack of a simplified version of community based Monitoring and Evaluation process has been as one of main challenges.

In response to the challenges, the Lake Victoria Environmental Management Project (LVEMP) designed a simplified M&E tool that could be used at community levels. Community M&E Registry Book – is a simplified manual tool that requires recording joint and agreed observations between visitors/technicians and Community leadership and Management; it records technical advices for the community understanding; it capture key events/issues happening in the community so as to ease implementations and follow – ups, it provide directives and agreed list-of actions to be done.

This paper wishes to share the usefulness of this tool in recording joint observations, gathering information, sharing and dissemination. The system enables effective collection, organization, storage, retrieval, analysis and presentation of a wide range of community data and information including, implementation progress. The tool can be used to assess progress as well as instilling sense of basic self participatory M&E skills among community members, transparency, knowledge building, empowerment and thus improved leadership and good Governance at community level.

Impact on Policy though Science Journalism Capacity Building – Evaluating a Mentoring Programme for Reporters in Africa and the Middle East

J. Lubinski1, M. Schanne2, C. Spurk2
1 World Federation of Science Journalists (WFSJ), Monitoring Evaluation Learning, Bonn, Germany
2 Zurich University of Applied Sciences, IAM Institute of Applied Media Studies, Winterthur, Switzerland

The capacity building of science journalists in developing countries has so far been mostly limited to short term training programmes in which reporters and editors of news media are trained to report on health, environment, science or technology. Evaluations of these trainings were usually very limited, if at all existent. In the past years the SJCOOP-project for science journalists in Africa and the Middle East has been developed in which young journalists receive support from experienced science journalists via a three-year mentoring process.

As evaluation framework a combination of two methodologies is used: Outcome Mapping and the Logical Framework Analysis. This joint approach brings the opportunity of evaluating results and the process of this intervention in a flexible, participatory and yet scientific way, based on a testable theory of change. The aim of this new framework in media development is to achieve a better evaluation methodology for journalism/media projects as well as to reach an understanding of the various effects of the interventions and their reasons, i.e. how and in which environment better science journalism can actively help government and donors’ efforts in achieving the Millenium Development Goals, for example by helping to improve health services, by breaking public taboos or by holding governments accountable.

The evaluation assesses the quality in science reporting, the relations and the process between mentors and mentees (and between mentees and scientists), the development of journalists’ national associations, the networks of science journalists (and scientists) emerging though SJCOOP and the impact of specific stories/articles on the public agenda and policy change. For the latter first cases are presented and a taxonomy of different agenda setting cases is discussed. The results may lead to new approaches in journalism training, media development and science communications.

Keywords: Mentoring; Outcome mapping; Agenda setting; Media development; Science journalism
Standardizing Cost-Effectiveness Results of a Social Service Program: a Clustered Sampling Analysis

L. Vilarins¹, M. Fontes¹, R. Crivelaro¹, F. Pereira¹, E. Lopes²
¹Social Services of Industry – Brazil, Unitep, Brasília, Brazil
²John Snow Brazil Consultancy, Research and Social Impact Evaluation, Brasília, Brazil

Background
As in the case of the Human Development Index, which standardizes rates of life expectancy, per capita income, and education level from all countries in the world, multi-site evaluation results may also require the use of a standardization technique. Standardization facilitates the identification of specific patterns and reduces variations across similar variables. In the case of cost-effectiveness analysis, high variations in cost and effectiveness rates may distort final results from multiple geographical areas.

Objective
In 2009, the Brazilian Industrial Social Services (SESI) implemented the Citizenship Rights Event (Ação Global Program), offering 1.3 million individuals access to basic social and health services. The main objective of this study is to produce a cost-effectiveness analysis as part of an impact evaluation of the program clustered around all 27 states of Brazil.

Methods:
Based on objectives, goals and indicators established in a logical framework of the program, a uniform 4 steps cost-effectiveness analysis was conducted across all 27 Brazilian states (clustered sample, n = 300). The 4 steps were: a) identifying financial per capita cost in each state; b) calculating the proportion of 41 efficacy variables; c) defining rates of efficiency and; d) standardizing a multi-site cost-effectiveness index.

Results
Per capita costs and rates of efficacy showed significant different ranges across states. For example, in the state of Paraíba, per capita cost of the program was 13 times higher if compared to the state of Piauí. However, using the cost distribution for all 27 states and the use of a standardizing multi-site cost-effectiveness index fit a normal curve. The final normalization test (swilk test) reached p-value > 0.05, confirming the null hypothesis of normal fit. This was also the case for final efficacy rates across all states.

Conclusion
The use of a standardizing multi-site cost-effectiveness index formula promotes better comparison of results across geographically clustered samples.

Keywords: Cost-Effectiveness; Citizenship Rights; Impact evaluation
R&D Evaluation: Regional, National, International

R&D evaluation in Italy. Moving towards a more institutionalised system.

G. Sirilli1, S. Giuffrida1, F. Tuzi1, A. Silvani1
1ISSIRFA – CNR, ISSIRFA, Roma, Italy
2National Research Council, DG Office, Roma, Italy
3National Research Council, DG, Roma, Italy
4DG – University of Milan, DG, Milan, Italy

R&D evaluation in the public sector in Italy is undergoing a deep transformation. In the 1990's R&D evaluation was not sufficiently widespread in the country and the European “push” played a significant role – in particular as far as methodologies are concerned. In recent years various evaluation exercises have been carried out on the basis of formalised initiatives. These exercises covered various aspects of the R&D and innovation endeavor: researchers’ scientific productivity (including national comparative analyses), regional innovation policies (evaluations carried out both by the Regions and by independent evaluators) and efficiency and effectiveness of universities and public research institutions.

The paper addressed the issue from various angles:
- the evolution of the so called “institutionalization of evaluation”, in particular with reference to the role played by the various actors in the evaluation process, and the methodologies used,
- to what extent the data availability influences the design of evaluation exercises, in particular the demand of evaluation from policy makers,
- how the national exercises compare with similar ones carried out in other countries, in particular if Italian national features may be identified which justify original evaluation models,
- the impact of R&D evaluation in terms of policy making, in particular in orienting the allocation of resources.

The four main issues mentioned above are dealt with both in the general national context and with reference to the evaluation of the scientific infrastructure of the National Research Council, the largest national public research agency.

Overall, the impact of R&D evaluation in Italy has been rather limited and the prospects for the future, given the lack of resources and the stickiness of the system, do not look so good.

Keywords: R&D evaluation; Innovation; Public policies; Regional development

Evaluation of Research and Development activities – the Risk Sharing Finance Facility – a joint European Union/European Investment Bank financed instrument

W. Schmidt
European Investment Bank, Operations Evaluation, Luxembourg, Luxembourg

Private investment in research, development and innovation has been and still is below the level necessary to achieve the goals set by the Lisbon agenda. In addition to the lack of private investment, capital market finance is scarce due to the high risk and uncertainty associated with R&D investments. The financial crisis has further aggravated this situation.

In order to foster private sector RDI investment the European Union, the European Commission and the European Investment Bank (EIB) established in 2007 a new financing instrument, the Risk-Sharing Finance Facility (RSFF), to improve access to loan finance (contribution being maximum EUR 1 bn each until 2013 – initially half has been committed). The RSFF is a debt-based facility to create additional financing capacity of up to EUR 10 bn in support of eligible RDI activities.

This EES paper will outline the scope and methodological approach used for this private sector evaluation of research and development activities within the European Union. It covers both project specific evaluation results, as well as an analysis of the entire portfolio of RSFF operations at programme level. It will build further awareness for research and development evaluations by highlighting the main challenges and complexity dimensions of these evaluations. It will discuss the main impacts and lessons learnt from the evaluation together with the results of the first independent evaluation of this joint facility.

Keywords: R&D evaluation
Evaluation Science: A vision for evaluation in public interest

Evaluation Science: A Vision for Evaluation in the Public Interest

S. Donaldson
Claremont Graduate University, Behavioral & Organizational Sciences, Claremont, USA

Does evaluation deserve the status of a “science”? How do we best incorporate scientific literacy into evaluation practice? This roundtable will explore a new integrative perspective on how best to practice evaluation in the era of scientific and evidence-based decision making. Participants will be presented with and asked to discuss some of the latest thinking about what counts as sound scientific evidence in evaluation practice, scientifically sound evaluation methods, evaluating complex interventions and theories of change, and how to integrate scientific research and evaluation. Is a serious move toward requiring more rigorous scientifically-based evaluations in the public interest?
The role of Non-Scientific Knowledge in Evaluation

The Role of Non-Scientific Knowledge in Evaluation

U. Schiefer1, A. Caetano1, A. Carvalho3, L. Bal-Doebel4

1Iscte IUL Avenida das Forcas Armadas, Sociology, Lisbon, Portugal
2Mai, Sg, Lisboa, Portugal
3Cea Iscte IUL Avenida das Forcas Armadas, Sociology, Lisboa, Portugal
4EP, Romania, Bukarest, Romania

Evaluation, as well as planning, strives to use scientific knowledge as a base for its findings: knowledge produced according to, fiercely debated, internationally recognised. The success of science and technology which lead to the paradigm of the “knowledge society” adds legitimacy to the scientifically-based evaluation process. A good part of evaluation research is dedicated to the methodology of the production of scientific knowledge.

Societies produce many different strands of knowledge and scientific knowledge is only one of them. Evaluation, although recognising and stressing the heuristically correct knowledge production, feeds on many strands of knowledge of evaluation stakeholders from fields very far apart, like political, strategic and operational, social and cultural, artistic, spiritual and, last but not least, practical knowledge.

The ambivalence of cultures where “scientific evaluation” and “scientific decision-making” are not well established incorporates the search for pre-reflective knowledge such as found in religion, superstition, traditional knowledge as transmitted through oral tradition, common sense, time-honoured rules and regulations, custom, philosophical wisdom, etc. Even in cultures where “scientific decision-making” is promoted, there is evidence of subcultures where reflection is based less on scientific knowledge and more on other types of knowledge.

On the level of knowledge and knowledge production, experience – which might in some cases even result in forms of wisdom – plays a very important role, even if is only indirectly acknowledged. On the level of perception, many times intuition and other forms of subtle knowledge production come into play.

Most of the (non-scientific) context knowledge tends to be expunged from the final evaluation results which have to conform to the standard of applied science in order to be taken seriously by decision makers and other evaluation users.

We argue that this knowledge deserves to be included in evaluations – how?

Keywords: Non-Scientific Knowledge Strands; Perception and Intuition; Heuristics and Knowledge Production; Knowledge Validation
Developing capability

Professionalizing Evaluation: Is Accreditation the Next Step?

R. Picciotto¹, S. Menon¹, L. Morra-Imas², J. Quesnel³, H. Simons⁵, F. Jordan⁶
¹ EES (Chair)
² UNDP
³ IDEAS
⁴ IOCE
⁵ UKES
⁶ CES

Background
As a discipline, evaluation has developed its own distinctive identity and a core set of principles as well as wide range of tools and methods. Towards enhanced quality of evaluations, national and regional evaluation associations have published evaluation guidelines. Since then, their focus has shifted to the identification of the knowledge, skills and dispositions needed to live up to good standards of evaluation practice.

Objectives
The proposed panel will debate what lies ahead in the professionalization of evaluation. It will outline the diverse initiatives underway at national, regional and global levels and examine the pros and cons of proceeding with accreditation of evaluation training programs and/or the adoption of formal professional designation processes.

Methods
Panel members will elicit audience participation by articulating the positions taken by various evaluation associations towards the issue of professionalization and presenting contrasting views about the case for and against the traditional measures taken by the long standing, self-governing professions (e.g. lawyers, accountants, engineers and physicians) to verify that adequate education and training standards are met and that only registered individuals are allowed to carry a professional designation label.

Results
The session will be facilitated in order to contribute to (i) a better understanding of the issues raised by evaluation professionalization; and (ii) improved coherence of the efforts currently underway worldwide towards the goal of enhanced evaluation capabilities and improved evaluation practice.
Policy evaluation and rural development

Impact assessment of the Flemish rural development programme – How to focus on main effects?

R. Dessers
IDEA Consult, RSO, Brussels, Belgium

IDEA Consult has been contracted for the ex-post evaluation of the Flemish Rural Development Programme 2000–2006 and the ex-ante and Mid-Term of RDP 2007–2013. DG AGRI proposes 7 impact indicators, which is insufficient to answer over 150 evaluation questions. IDEA Consult (1) collects the core themes occurring in the evaluation questions and lists all the possible links between the different measures of the RDP and those core themes, in order to capture the overall programme effects. Due to evaluation budget limitations, IDEA Consult afterwards (2) focuses on those measures which are expected to have a high impact on the core impact themes.

1. Collection of core impact themes

Out of 150 evaluation questions, the evaluator collects a list of key words, i.e. the core impact themes/objectives. The various measures of the RDP are analysed on their (partial) contribution to these core themes, in order to capture the overall programme effects.

12 core themes are classified into 4 thematic components:

1. Competitiveness of agriculture and agrofood sector
2. Environment, nature and the landscape
3. Rural services & infrastructure and diversification of rural economy
4. Governance and convergence of rural areas

2. Scope

Due to the limited budget for evaluation of the Flemish RDP, only the relevant links between measures and impact themes are profoundly analyzed. A link between a certain measure and an impact theme is considered relevant if the measure is expected to have a high impact on the impact theme. The score on ‘expected impact’ is based on an interactive workshop with beneficiaries, the available output-indicators and budget of each measure and existing causal evidence (from – not necessarily RDP-related – empirical research). Based on this ‘expected impact’-criterium, (combinations of) measures and impact themes are selected (marked in grey in figure above) for which impact assessment is carried out.

Keywords: Rural development; Flemish; Scoping
Evaluating the improvement of quality of life in rural areas

S. Cristiano, S. Tarangioi
INEA – National Institute for Agricultural Economics, Rural policy, Rome, Italy


In the framework of the European Common Agricultural Policy, the enhancement of the quality of life in rural areas is one of the major strategic objectives to be addressed by a menu of measures within the Rural Development Programmes 2007–2013.

The evaluations are expected to assess the improvement of quality of life in rural areas as effect of programmes’ implementation.

The study is motivated by the need for developing an appropriate evaluation method that leads to gathering meaningful information for a broader understanding of the quality of life in rural areas, including the subjective well-being’s dimensions and its determinants and feeds the policy designs on this specific domain.

Selections of some current literature on the multidimensional nature of quality of life have been used as conceptual basis for analysing the extent to which the European evaluation framework for rural development programmes (EC 1999, 2006) – based on the intervention logic model, the use of economic indicators and evaluative questions – is able to capture the relevant dimensions of well-being rural people’s lives.

A part of the research is based on the analysis of ex-post evaluations carried out in Italy.

The paper proposes a wider integrated evaluation approach to be used in the context of the evaluation of impacts of rural development programmes, that through the combined utilization of quantitative and qualitative indicators and additional evaluative questions, allows a more comprehensive assessment of quality of life in rural areas.

Keywords: Assessing quality of life; Measuring well-being; Evaluation; Qualitative indicators

Evaluating the local governance capability: learning from Leader approach

M. Bolli, B. Zanetti
INEA – National Institute for Agricultural Economics, Structural and territorial analysis and services for agricultural development, Rome, Italy

New trends orienting rural development policies affirm the importance of a territorial approach in order to stimulate integrated development processes of rural areas, characterised by multiple dimensions: economic, environmental and social ones. This implies that on rural areas necessarily have to operate different policy instruments and actors and that the impacts come from a combination of them.

A territorial perspective is ever more topical for rural policy because the European countries, even if with differentiation, are tending toward an administrative decentralisation in favour of regions and sub-regional level of government, enriching in this way the institutional framework of rural policy.

Local governance and specific programming and implementing mechanisms, represent therefore central themes for evaluating the functioning of rural policies and improving the overall efficiency and effectiveness thereof.

This has important implications for the evaluation of RDP and utilization of the results, putting in evidence the need of evaluative analyses which focus much more on processes, rather than only on mere quantitative results. Some studies carried out over the last years such as the European Project RuDI, that aims at improving the understanding of Rural Development Policies (RDP) impacts and, specifically, the case study we illustrate, concerning the experience of an Italian Local Action Group, show the importance of assuming a territorial and integrated approach to evaluation. In order to examine rural policy ‘process-effects’ that can be identified as pathways through which to influence the future impacts of policies, the study mainly focused on relevant aspects of the process of design and implementation of a local strategy of development: drivers of changes, procedures, innovative solutions in terms of partnership working arrangements and mechanisms of coordination at different levels, involvement of local stakeholders, integration of measures and different policies.

Keywords: Territorial approach; Processes; Integration

Case-study findings: Implementation of the measures supporting diversification in rural area in the Vysocina region

M. Konecna 1, E. Kozakova 1, J. Prazan 2, T. Ratinger 1, Z. Bednarikova 1
1 UZEI – Institute of Agricultural Economics and Information, Regional and Rural Development, Prague, Czech Republic

In the 7FP project RuDI – “Assessing the impact of rural development policies (including Leader)” a case study was carried out to answer the question: “What are the obstacles to the effective implementation of measures 311, 312 and 313, especially regarding to the farmers’ capacities to apply for the support?”

This paper describes findings and contextual factors which had impact on the effectiveness of the implementation of measures supporting the development activities in rural area with further focus on the evaluation of the social and human potential of region and setting of criteria of accessibility for the applicants. The respective theoretical base for capturing the barriers to the more larger uptake of diversification support measures followed form the institutional economics, particularly, the theory of transaction cost, theory of prisoner’s dilemma and social capital theories as well as from regional development theories.

Semi-structured interviews with farmers and representatives of the administration (the Ministry of Agriculture, Paying Agency) and support organisations as agricultural agencies or advisors, etc. were the core of the case studies. Further, statistical data and data from previous research and relevant literature were used to complete the picture.

The main findings which emerged from the interviews indicate the low level of the common understanding of rules, low capacity of applicants to accept frequent changes in the conditions and rules of the measures and insufficient targeting of the measures lacking link to the regional potential.

The findings of the project indicate that institutional factors are important for the success of policies and thus they ought to be considered in evaluation of policies as well as in policy planning and design (referencing to the coming new programming period 2014–2020, in particular).

Keywords: Rural development program; Social and human capital
Influences on evaluation Processes and Results

On going evaluation of rural development programmes in Italy: steering group as pathway of influence

V. Fucilli, S. Cristiano, A. Monteleone

University of Bari, Progesa, Bari, Italy
Istituto Nazionale di Economia Agraria, Inea, Roma, Italy

Evaluation system of the European Union rural development policy has evolved considerably compared with the previous programming cycle. In line with the more strategic approach of the implementation methods of rural development policy, the new evaluation system has introduced changes in comparison with the past. Among these changes, there is the organization of on-going evaluation whose activities are realised throughout the programming period.

Innovations and complexity of evaluations of rural development programmes (RDPs) need the definition of structures able to manage processes and foster use of results. In this specific context it appears particularly important to understand what kind of structures have been set up, by comparing various experiences in Italy and considering that EU confirms the essential function of Steering Group “in order to ensure a high quality of the evaluation” (EU, 2006, p.8).

The aim of the paper is to analyse the current setting up of those structures, their roles and functions, models and configurations, as a way of improving evaluation influence (Kirkhart, 2000). It is an attempt to investigate those structures that have been created to influence the quality of evaluations. The research, in fact, originates from the observation that, although there is undoubtedly a mounting interest on better structuring evaluation use, there does not appear to be a systematic analysis of this specific subject, in particular in the domain of the RDPs evaluations. Henry and Mark (2003) and Mark and Henry (2004) offer the conceptual foundation for this study that has been conducted on the basis of the contents of Managing Authorities decisions on Steering Group (or other structures), Rural Development Programmes evaluative plans, call of tender documents, annual progress reports, telephone interviews and questionnaire.

Keywords: Evaluation; Rural Development Policy; Influence; Steering groups

How the political context affects evaluation methodology

T. Azzam
Claremond Graduate University

A simulation study was conducted in an attempt to examine how evaluators modify their evaluation design in response to differing stakeholder groups. In this study evaluators were provided with a fictitious description of a school-based program. They were then asked to design an evaluation of the program. After the evaluation design decisions were made, evaluators were presented with feedback from three differing stakeholder groups (i.e. decision maker, implementer, recipient) either endorsing or rejecting the evaluation design. Evaluators were then given the opportunity to modify (or not modify) their original design in response to stakeholder feedback. The findings revealed that the more political power or influence stakeholder groups held over evaluation logistical factors (i.e. funding, data access), the more evaluators were willing to modify their design choices to accommodate perceived stakeholder concerns. These design modifications were typically implemented to ensure data access, reduce stakeholder resistance, and increase stakeholder buy-in.

Keywords: Politics; Stakeholders; Methodology; Simulation
To remain “independent” on a small evaluation market – challenges and choices

P. Teixeira, S. Monteiro
Logframe – Consultoria e Formação Lda, Evaluation Unit, Lisboa, Portugal

The paper is a reflection on both the theoretical and practices levels on how to promote and develop independent evaluation practices on a small market (in this case Portugal).

The reflections is done on several levels that range from the real possibility to do evaluation without “contamination” from other phases of the policy, programme or project design or implementation to how the small market with low budgets is not attractive to bigger (even international) players and because of that it stays confined to the same national players. In fact both of the above mentioned conditions are well connected because it’s almost impossible completely avoid in a market like Portugal that stakeholders involved in the design or implementation phases for policies, programmes or projects not to be involved on the evaluation processes and that’s why we talk about “contamination” and a less clearer system.

In a context like this evaluators and Administration bodies face several challenges, like:

• The correct and clear definition of the evaluator role(s) and it’s professional mandate;
• The valorization of evaluation as a credible information source for the (re)design and/or revision of policies and/or programmes;
• The effective recognition, by the policy makers and others, of the usability of evaluation results;
• The need for professional structures that defined the evaluation profession in Portugal (we have a small and new evaluation association that could play a role here);

These challenges could be more easily surpassed if the EU programmes, because they were the real catalysts of the evaluation practices in Portugal, were more demanding, at the national level, on the quality control of the evaluations performed, had more realistic budgets (if you don’t realistically budget something is because you don’t really think it’s worthy) and on the effective separation of roles an ethic on evaluation practices.

Keywords: Ethics; Roles; Independence; Credibility; Evaluation value
Counterfactual analysis

Assessing impacts in realworld evaluations: practical alternatives to the conventional statistical counterfactual

J. Rugh
RealWorld Evaluation, Independent International Evaluator, Sevierville TN, USA

There is increasing pressure by major donors of international development programs for “rigorous impact evaluation” of the types of programs they fund. The assumption seems to be that these require evaluation designs that assess program impacts through experimental and strong quasi-experimental designs, including randomized control trials (RCTs). These research designs use statistically matched control groups to define the counterfactual condition — What changes would there have been in the condition of the program population if the intervention had not taken place? However, this approach largely ignores the fact that these strong quantitative designs can only be applied in a small proportion of realworld program evaluations, and very little advice is offered on how to assess the extent to which the observed changes in outcomes or results for the majority of realworld development programs can be attributed to the program intervention and not to factors unrelated to the program.

While claims are often made concerning the availability of alternative ways to define the counterfactual, it is difficult to find specific examples of how these alternative methodologies have been applied and interpreted in the field; and even more difficult to locate discussions of the methodological validity of findings and conclusions from these alternative approaches.

The proposed roundtable discussion will build on a successful think tank during the 2009 AEA conference where more than 50 participants shared experiences on a wide range of quantitative, mixed method and theory-based approaches used to define and assess rival explanations of the observed changes in the project population. Participants in this roundtable will have access to a document was produced describing more than 30 alternative approaches. Participants will be invited to share their own experiences and to identify additional approaches for selecting appropriate comparisons and applying alternative forms of the counterfactuals.

Keywords: Counterfactual
Performance evaluation of a cervical cancer screening program: an integrated model for local analysis at Rio de Janeiro/Brazil

T. Oliveira, E. Santos, M. Cruz
Oswaldo Cruz Foundation, National School of Public Health, Rio de Janeiro, Brazil

Background
Despite having high chances of cure when early diagnosed Cervical Cancer is still an important public health issue. Worldwide national or local control programs face a number of challenges in order to achieve lower incidence and mortality rates, especially at developing countries or among specific population groups. Brazilian official data show lower related mortality rates since the implementation of an organized national control program but process performance evaluation is needed to highlight potential improvements.

Objectives
To present a model for evaluate the performance of a Cervical Cancer control program and its applicability on a local context.

Methods
Based upon an extended bibliographic review and considering the theoretical conceptualization by Contandriopoulos et al., an integrated model for performance evaluation of cervical cancer control practices has been developed and applied on a testing base as a case study on a local program at Rio de Janeiro/Brazil. Data were collected by interviews with program managers and health professionals, process observation and analysis of program epidemiological data.

Results/Application to practice
Evidences show that although guidelines and indicators for performance monitoring are reasonably well established, health professionals and program managers need evidences of how individual, programmatic, and social vulnerability factors influence local programs performance. The introduction of innovative technologies for prevention as HPV vaccines should consider those factors on particular local contexts. In Rio de Janeiro the need of implementation of new strategies for enhancing program performance where found specially for technical components related to social communication and health professionals capacity building on monitoring and evaluation. Networking organization of health services should also be considered for enhancing access, coverage and adherence of target women to systematic attendance of screening based programs.

Keywords: Performance evaluation; Cervical cancer control

Defining performance indicators towards a participative evaluation in high education. The case of an University training course

M. Freddano, A. Siri
University of Genoa, DISA – Dipartimento di Scienze Antropologiche, Genova, Italy

Students need to have a clear understanding of the learning objectives and criteria against which their achievement will be judged. Self-, peer-and collaborative evaluation are designed to improve leaning quality and active engagement of students in evaluation and knowledge building. They suggest reflection, comparison and discussion based on student cooperation -and in formulating opinions on their activities and/or that of other students in relation to clear defined evaluative criteria. Students can preliminary be involved in negotiating their final examination criteria.

This study aims to illustrate the case of an University training course that has proposed a learning assessment model such as a genuine participative process and where students participated to the definition of the final examination criteria using face to face and online tools of social research. They have been engaged in an evaluative brainstorming moderated by the teacher discussing on and selecting, the dimensions and the indicators of the final student examination. Then an individual and online activity has been done using a forum to facilitate students in selecting some of the individuated indicators, putting them in a descending order of importance and giving them a weight by an auto-anchored scale.

The indicators have been defined by calculating the average of their rankings and their weights. The results have been published online and shovewd to the students during a face to face lesson by tables and graphs. The aim was to start up a participated activity to validate the criteria and indicators and to select the modality of their implementation.

We have organized two panels: the “teacher panel” composed by the members of the Institutional Examination Commission and the “student panel” composed by rotating students.

This participative experience let students know how the evaluation is a comparative as well as a complex activity.

Keywords: Self-evaluation; Peer-evaluation; Collaborative evaluation; High education evaluation; Indicators
How can performance measurement lead to improved performance in public sector organizations: The role of contextual conditions

S. Goh
University of Ottawa, Telfer School of Management, Ottawa Ontario, Canada

Performance measurement and performance management have become the focus of central agencies in government worldwide to address the issues of accountability and transparency. The underlying assumption is that government agencies and departments are viewed traditionally as inefficient and under performing and there is increasing public scrutiny of how governments deliver services (Kelman, 2006).

Increasingly, however, scholars have been studying and debating about its efficacy in achieving the objective of improving the performance of public sector organizations (Radin & McGuire, 2004, Radin, 2006). The current evidence is that performance management and measurement faces many challenges and obstacles in implementation. In fact, studies have shown that there are significant problems and questionable benefits of performance management to the public sector. We argue that these problems are due to the neglect and lack of focus on context and the management of performance management.

Most discussions of performance management in the public sector ignore the contextual conditions that can impact on the effectiveness of performance management systems. As Lye (2004) puts it, “in designing performance measurement systems researchers have failed to comprehensively examine the rich interdependencies between contextual factors and the use of performance measures”. In this paper we will discuss the implications of three contextual factors, managerial discretion, organizational culture and enabling formalization and its impact on performance management systems and the potential gains to improving organizational performance. We will develop a theoretical framework that links these three contextual factors to a performance management system to support our assertion of the importance of these contextual variables in explaining and predicting how performance measurement can lead to improved performance in public sector organizations.

The paper will conclude with a discussion on some practical implementation issues and how to overcome some of the environmental and institutional constraints in the public sector.

Keywords: Performance management; Performance measurement; Contextual conditions; Managerial discretion; Culture enabling Formalization

Implementing Results Based Management (RBM) in the public sector: Lessons from the field for Evaluators’

R. Lahey
REL Solutions Inc., President, Ottawa, Canada

Background

Results Based Management (RBM) or ‘managing for results’ has been the driving management model in the public sector for many countries through most of the 2000s. Moving from the theory to the practicalities though is generally not an easy or quick step.

The expectations about RBM have been raised very high (sometimes unrealistically high) by many organizations, central agencies, public sector theorists, etc. and the gulf between the concept and the reality is often quite wide.

Important lessons for Evaluators can be drawn from the experience, both good and bad, of other organizations and countries.

Objective

This paper presents some recent international experience of the author in helping organizations across the globe to develop and put in place management approaches that rely on results-based monitoring and evaluation systems. A focus will be on the lessons learned about the practical realities in developing and implementing RBM in public sector organizations and the role that Evaluators in any country can/should play in supporting RBM implementation.

In speaking about the practical realities in moving from the concepts of results-based M&E to actually implementing workable systems in organizations/countries, the presentation will address ‘the good, the bad and the ugly’ in implementing the concepts of RBM.

The presentation will draw on the author’s experience in working in three very different venues: a large NGO in Canada; a UN organization; and a developing country. In all three cases, there was considerable interest in developing and implementing the concept of results-based management and a results-based M&E system. All three cases exhibited a bit of ‘the good’; ‘the bad’ and indeed, ‘the ugly’. The intent will be to draw broad ‘lessons learned’ and advice/guidance for Evaluators regarding the role that they can/should play.

The presentation will have a broad format comprised of three parts: (i) the context and identification of the cases that are serving as the basis for discussion; (ii) ‘Lessons learned’ about the practical realities in developing results-based M&E systems and implementing RBM in public sector organizations; and (iii) role(s) that Evaluators can/should play in supporting RBM implementation.

Keywords: Evaluation; Results based management; Performance monitoring; RBM; International
Evaluation of innovation policies and programmes

New approaches to analysing key determinants of National Innovation Systems in developing and transition economies

L. Müller¹, G. Meier zu Köcker²
¹VDI/VDE Innovation + Technik GmbH, International Technology Cooperation and Clusters, Berlin, Germany
²Institute for Innovation and Technology iit, Berlin, Germany

When interviewing policy makers in emerging and developing countries about possibilities for stimulating innovation within their economy, it often becomes apparent that they look for structured descriptions of National Innovation Systems (NIS) and clear recommendations for improving the performance of their specific NIS. Rather than asking for scientific models of the functionality of an NIS or for sophisticated, statistically based performance indicators they are more interested in the practical assessment of their economic environment as regards to innovation.

Indeed, some existing analyses lack precise recommendations on how to optimise the effects of innovation capacities particularly in an environment where the resources for public investment are little. It is therefore crucial to identify those determinants of an NIS, which can be improved with the tools and financial means available for third party donors.

Our method “Analysis of National Innovation Systems – ANIS” provides a comprehensive examination and evaluation of the status of existing national innovation systems. It is in line with the new approach of indicator-based studies relying on quantitative data generated by the evaluation of expert interviews. The concept differs from traditional benchmarking studies on innovation performance, such as the European Innovation Scoreboard.

The ANIS approach is based on the assumption that an NIS is mainly influenced, at national level, by 30 determinants. ANIS takes up this challenge by providing an indicator-based assessment of these determinants, each of which reflects an aspect of the innovation system. The determinants may be grouped according to a three level hierarchy:

- Macro Level: Innovation Policy Level
- Meso Level: Institutional Innovation Support Level + Programmatic Innovation Support Level
- Micro Level: Innovation Capacity Level

A comparison between the determinants of these different levels allows the identification of key policy areas requiring a potential intervention to strengthen the NIS.

Keywords: Analysis; National; Innovation; Developing; Determinants

Gender Evaluation Methodology for Internet and Information and Communication Technologies (ICTs): A Practice for Policy Implementation

L. Simerska, C. G. Ramilo
Association for Progressive Communications, Women’s Networking Support Programme, Manila, Philippines

Gender Evaluation Methodology for Internet and Information and Communications Technologies (GEM) is an evaluation tool developed by The Association for Progressive Communications Women’s Networking Support Programme (APC WNSP). GEM integrates a gender analysis into evaluations of initiatives that use Information and Communication Technologies (ICTs) for social change, development, and poverty eradication. Since 2003 it has been used by ICT initiatives, project managers and project staff using ICT in projects without a specific gender or women’s focus, evaluators working in the IT field, donors and development agency staff working in the IT field, gender focal points that support women’s and IT issues, policy makers, ICT planners, and consultants in the area of gender and ICTs. GEM has been applied in a number of ways all over the developing world and adapted to specific areas of use such as: Rural ICT Development, ICT Policy Advocacy, Telecentres, Localisation. GEM also maintains its network of users, both individuals and organisations, called GEM Practitioners Network (http://www.genderevaluation.net/mygem/).

The focus of the this paper is to present examples of GEM’s application where evaluation brought about successful policy implementation and changes in public interest. Participatory evaluation process that GEM encourages, determines whether ICTs are really improving women’s lives and gender relations as well as promoting positive change at the individual, institutional, community and broader social levels.

This has been the case of the three presented evaluation studies: evaluation focused on discrimination against women in the e-governance scheme in rural India, evaluation study of the use of telecentres run by Philippine Community eCentre Network, a multi-stake-holder learning and collaborative community for eCentres in the Phillipines, and the evaluation process at telecentres in southwestern Colombia in partnership with academics from Universidad Autónoma de Occidente.

Keywords: Gender; Information and communication technologies; Development; Participatory evaluation process; Social change
Evaluating legislation

Presentation of the main conceptual and methodological challenges when evaluating non pending programmes

E. Monnier1, J. Grear2, K. Williams3

1 Euréval, Chairman, Lyon, France
2 European Commission, DG INFSO Evaluation Unit, Bruxelles, Belgium
3 Organisation for Economic Cooperation OECD, Council Secretariat, Paris, France

Background

For more than 20 years the European Commission has been a leader in disseminating evaluation and evaluation practices. Evaluation has been mostly limited to spending instruments though, when most of the European policy-making is actually law-making. In fact and according to estimations, the Commission implements some 200 spending programmes, and 2,500 legislative or soft-law instruments that constitutes most of the influence of the European Union on citizens, governments and businesses.

Objectives

After the introduction in 2005 of mandatory impact assessments before European regulatory instruments can be adopted, it is only a question of time before European laws start to be evaluated just as spending programmes have always been.

DG MARKT pioneered the domain in 2007 by issuing a first guide on how to evaluate legislation and other non-expenditure instruments. DG INFSO has engaged in producing its own guide in 2010. The guide explores the factors of complexity that are inherent to the evaluation of legislation and put it apart from standard evaluation: the number of potential instruments, the complexity and unpredictability of the causality chains, the number of stakeholders at stake, the uncertainty caused by the transposition process and the chances of unexpected impacts, among others.

Keywords: EU guideline; Evaluating legislation; Conceptual framework; Intervention logic; Tools and method
Theories and visions I

A question of theory or implementation? The role of assumptions in evaluating global agendas for aid effectiveness.

E. Lithman
Former Head of the Evaluation Department of the Swedish International Development Cooperation Agency (SIDA), Europe, Stockholm, Sweden

Background
The Paris Declaration and the Accra Agenda for Action are examples of global agendas that define principles and actions intended to increase the effectiveness of development aid. How these agendas and their implementation should be evaluated is the focus of an emerging professional discussion, e.g. the animated discussion hosted by IDEAS during the month of April 2010 on the evaluation of the Paris Declaration. Attempts have also been made to evaluate the principles for EU development cooperation (the so-called 3 C evaluations on the principles in the Maastricht treaty) as well as aid effectiveness at country level.

Objectives
The objective of the paper is to illustrate and discuss the assumptions, scope and approaches used when evaluating global aid effectiveness agendas. In addition the paper will discuss the potential as well as the limitations of these evaluations to inform policy and practice.

Method
A number of evaluations are selected and analyzed for underlying assumptions. The approaches such as theoretical frameworks and the methods employed will be described and assessed.

Conclusions
Rather than presenting conclusions the paper will contribute to the discussion regarding the role of evaluation in advancing the understanding of how global agendas influence the effectiveness of aid.

Keywords: Aid effectiveness; Paris declaration

Evaluable Assessment 2.0 – Or: On a Reflexive Approach to Evaluation Systems in the Knowledge Society

P. Dahler-Larsen
University of Southern Denmark, Department of Political Science and Public Management, Odense, Denmark

Evaluable assessment is a diagnostic and prescriptive tool which helps evaluators determine whether evaluation is appropriate based on a careful situation-specific analysis. Thus, evaluation is a situational rather than a universal good. Today, however, evaluable assessment is no longer particularly popular, but mandatory, comprehensive and repetitive evaluation systems are gaining ground supported by general social, political and managerial norms and values.

Can a form of evaluable assessment be re-vitalized in order to pave the way for a more modest, more reflexive, and more context-sensitive belief in evaluation? An updated version of evaluable assessment should take into account not only the complexities of evaluation systems but also the non-linear use of knowledge in contemporary society. The article offers a specific list of items in an updated version of evaluable assessment, and concludes with a discussion of the limitations of such approach.
Who is the evaluator? Perspectives on the changing role of the evaluator

I. Rydell, M. Wikner

1 Ramböll Management Consulting, Growth Analysis, Stockholm, Sweden
2 Ramböll Management Consulting, Welfare Analysis, Stockholm, Sweden

Background

In both literature and in the practical evaluation work, we experience a trend towards evaluation commissioners involving evaluators as integral to the entire policy process (see among others “From studies to streams: managing evaluative systems”, Rist & Stame, 2006). Traditionally, the role of the evaluator has been tied to the later steps of the policy process – e.g. retrospective evaluations. Today, the evaluator tends to be active in virtually all phases of the policy process; problem identification/needs assessment, recommendations for solutions, tailoring of interventions as well as implementation and evaluation of the same.

The role of the evaluator may thus be changing. The evaluator often has several functions in the same assignment: field expert, researcher, accountant, coach, etc. – as well as that of the “traditional evaluator”. It is our experience evaluators often lack a strategy to handle these new roles. Part of the problem is an uncertainty of how the new role should be understood, and what implications it has for the traditional role of the evaluator. If the evaluator does not understand the new role, or acknowledge the implications, there is a risk that we fail to expand the reach of evaluative knowledge according to these new demands. Therefore, there is a value in clarifying how the role of the evaluator should be understood and what implications this has on processes and methods.

Objectives

Our paper aims to present an overview of changes in the role of the evaluator as the assignment has moved from retrospective control to all steps in the policy process.

Method

Analysis of process and content of four reference assignments from the Swedish context.

Results

There is a need for a broadened understanding among both commissioners and practitioners of the role of the evaluator since this has implications on relevant evaluation processes, methods and tools.

Keywords: Role of evaluator; Broadened assignment
Partnership and policy

Caught in the paradigm? Exploring alternative approaches to evaluating service innovation

J. Moltesen¹, H. Noes Piester²
¹ Tavistock Institute, Research & Evaluation, London, United Kingdom
² Danish Technological Institute, Centre for Policy and Business Analysis, Copenhagen, Denmark

The paper explores how experimental service innovation projects in the public sector challenge standard government approaches to evaluating public sector initiatives and the challenges we encountered as evaluators of four service innovation projects launched by the Danish Government in 2009. Service innovation is considered a key transformative factor for businesses and the public sector and governments are currently exploring different approaches to promoting service innovation as a means to reducing the costs of service delivery whilst also increasing service quality. This presents a need to develop evaluation approaches which are able to assess the impacts of service innovation initiatives as well as their potential contribution to public sector transformation.

In the paper, we will analyse how a government can be caught in a paradigm in which evaluation is associated with the collection and analysis of quantitative data on output and outcomes such as overall number of patients, waiting lists and user satisfaction before and after implementation. We will argue that a continuation in this stream bears the risk of undershooting the potential impact of service innovation, because standard quantitative approaches disregard key aspects of service innovation like work and service culture, service design, interpersonal relations and organisational learning.

Based on our experiences from the evaluation of service innovation projects, we suggest that if governments want to reap the potential benefits of service innovation projects they need to be more open towards:

a) process evaluations based on qualitative data and
b) alternative data collection methods and techniques such as proof-of-concept and touch-point flow analyses. Finally, we will discuss how carefully constructed evaluations of service innovation projects can contribute to more large scale public sector transformation and systemic innovation.

Keywords: Service innovation; Evaluation methods; Public sector change

The challenges of the participation of Zambian Civil Society in Country-led governance evaluations

J. Njovu
Zambia Revenue Authority, Domestic Taxes, Lusaka, Zambia

Background
One of the key components of the Zambian Fifth National Development Plan is the governance programme. There are six components in the governance programme:

1. Administration of Justice
2. Constitutionalism
3. Democratisation
4. Human Rights
5. Accountability and Transparency
6. Coordination, Reporting and facilitation of Governance Initiatives

Zambia is mainly aspiring for local and joint ownership of evaluations and not foreign ones. It has been helped to localise by GTZ and Inwent Germany. Key stakeholders from Civil society and government participated in 2008 in the drafting of questions for inclusion in governance survey. The survey was carried out by the government’s Central Statistical Office. However, the government has not released the report. Civil society has not even seen the draft report.

Objectives
The objective of this presentation is to show the challenges that civil society faces in governance evaluations in Zambia.

Methods
The first governance survey was carried out by the government Central Statistical office using questionnaires and covered the whole country. There will be an evaluation of the performance of Members of Parliament during May 2010. This will be done by Governance Monitoring and Evaluation Committee.

Conclusion
Challenges of civil society:
- Central Statistics Office is viewed by many as lacking independence.
- Lack of experience in country-led assessments by Zambian civil society
- Lack of financial and technical resources to conduct country-led evaluations
- Lack of control on the output
- Challenges of reaching citizens and finding out their governance problems

There is need for independence of governance monitoring and evaluation (M & E) to be fostered by cooperating partners. Mentoring can be fostered between experienced M & E professional associations, like EES, and the Zambia Evaluation Association. This will enable Zambian civil society to have capacity to carry out independent evaluations.
Public policies evaluation process: The researcher situation within different evaluation mechanism

O. Heddebaut
The French National Institute for Transport research and Safety

This communication questions processes and practices when conducting public policies evaluation in transport and tourism domains. It examines different procedures implemented to face the legislative frameworks. Evaluation mechanisms are analysed for the State and the Nord Pas de Calais Region contracts in tourism domain, for the Lille urban travel plan evaluation (Heddebaut, 2000, 2008) and for large scale transport infrastructures evaluation (Kaufman et al, 2008).

Firstly the organisation of evaluation itself is interrogated as for the choice made to realize this evaluation between “an internal mode” by the own services of the evaluated structure or “an external mode” meaning exterior evaluators. Different technical or political committees composition and role they play and have in the whole procedure are analyzed.

Secondly, through these experiences it can be observed the posture and the influence of the exterior evaluator whether it be an expert or a researcher involved into these mechanisms settled to realize a given evaluation. Effectively, evaluators are confronted to antonymous positions described by Patton or Weiss: either they plead for involvement more or less intense within the evaluation approach for a better understanding of the real implementation. This contains the risk of collusion or adaptation (Azzam, 2010). Either they proclaim a pure independency and be as neutral as possible such as to keep their credibility. Moreover adopting the values defended by the evaluated policy is considered as a suspicious attitude because it suppose a diminution of their independency even if within evaluation domain exercise the “valuephobia” described by Scriven in 1983, seems to be not adapted to the evaluation exercise and even be illusory.

Thirdly, the way the evaluation results are used or not is tackled. Researchers hope their recommendation will be taken into account and be the foundation of further policies. Their consequences being new research objects.

Keywords: Researcher role; Independency versus involvement; Valuephobia; Neutrality versus integration and adaptation
Policy evaluation: Evaluation in Secondary Schools

Multi-method evaluation of the “hands-on” approach to science teaching in Italy

B. Romano
University of Genoa, Disa, Genova, Italy

With funding from a philanthropic foundation, a number of museums in northern Italy worked with elementary and middle schools to introduce an “hands-on” approach to science teaching. The ultimate goal of the program was to encourage students toward by exposing them to a more stimulating approach to science.

Evaluating the success of this type of education programme poses a number of challenges. First, the treatment is not one-shot, but it consists in a number of steps: the teacher must receive an initial training; then he/she must adopt the hand-on approach in his/her actual teaching practice; finally, the student must react positively to the innovation, which in turns depends on a number of contextual factors. The first task of the evaluation is to document this complex process, in order to identify issues and difficulties. One crucial issue concerns the sustainability of the program in those schools that first adopted it, after the support of the museum and of the philanthropic foundation is withdrawn.

A second challenge is represented by the choice of the outcome variable on which to measure the effect of the program. We chose the propensity on the part of the student to enroll in a high school scientific-oriented curriculum. We then adopt a counterfactual approach, attempting to establish whether the program made a difference in this respect.

We used a multi-method approach to answer the questions: surveys, in depth interview, focus group, secondary data analysis. In particular, the issues related to the implementation were explored using in-depth interviews with museum directors and questionnaires directed to the teachers. The issue of sustainability was investigated also by means of repeated interviews. Finally, the impact analysis was conducted by comparing the behavior of participating students with that of preceding cohorts of students taught by the same teachers.

Keywords: Counterfactual; Implementation; Sustainability; Professional development; Science education

Evaluation of pupils’ learning environments in upper secondary schools in Norway

L. Monsen
Lillehammer University College, Dep. of Social Science, Lillehammer, Norway – Lars.Monsen@hil.no

In this paper I will analyze data from a national pupil survey in Norway in 2009. The aim of the survey is getting inclusive data on learning environments in Norwegian schools with more than 300 000 thousand answers. In my analysis I will compare data from this national sample with a sample from one county – Hedmark county with answers from near 3000 pupils. The aim of this analysis is to find out the variation in learning environments between different schools in Hedmark county. From earlier surveys we knew that the variations between the 20 counties in Norway were rather small. In my analysis I found a rather astounding variation in pupils responses between different schools. I could identify school were pupils were more content with their leaning environments than in other schools. In the concluding part of the paper I discuss what we could do with this kind of information.

Keywords: Learning environments; Upper secondary schools; Norway; Electronic survey
Creative imagination in interactive evaluation

Creativity and empowerment in residential elderly care

V. Baur, T. Abma, Y. Voskes
VU University Medical Center, Medical Humanities, Amsterdam, Netherlands

In residential elderly care settings, older people live in an institutionalized context that can hamper feelings of independence and autonomy. Within traditional hierarchical relationships, professionals tend to fill in for residents what is good care and what are good services. Managers and caregivers have professional knowledge and moral intuitions about what contains good care. However, older people who live in a residential care home themselves have experiential knowledge about what is valuable to them. A recent development towards shifting power relations within the residential care institute is collective participation of residents in practice improvements through dialogue with managers, employees, and volunteers. This way, room is made for relational empowerment within the organization, a process in which everyone involved changes.

In this presentation we will describe our interactive evaluation of innovative ways of resident participation, using creative methods to include all voices and to transform negative group dynamics into positive. We will stage the Taste Buddies, a group of seven older ladies (age 85–94) living in a residential care home in the Netherlands. We will describe how they developed partnership relations with managers, employees and volunteers in improving the daily meals. Using creative methods (photo-voice, making a paste-up) turned out to be a useful instrument in shifting power relations and bringing forward the experiential knowledge and empowerment of people who used to feel powerless before. Moreover, also the interactive evaluation process itself is creative: the evaluator has to find creative ways of engagement that can only emerge gradually, in context and in partnership with all those involved.

Alternative format: ‘Interactive symposium’: 3 presenters giving a paper on the theme ‘evaluation and creativity’ and inviting the audience to participate in an interactive and creative session after the presentations.

Keywords: Creativity; Empowerment; Elderly care; Power relations
Developing ethical capability

Developing an approach for considering ethics in evaluation in the United Kingdom Government Department responsible for international development

K. Hinds
Department for International Development, Evaluation, East Kilbride, United Kingdom

Background
The UK’s Department for International Development (DFID) has a growing annual budget (currently almost £8 billion) and is increasing spending on evaluation. We have a centralised evaluation function commissioning evaluations externally, with evaluations also commissioned by our network of country offices, primarily in Africa and Asia. DFID’s 2009 evaluation policy committed us to producing common evaluation standards across the organisation, including on ethics. Our research department was also very interested in this issue. Existing guidelines relevant to the UK did not deal with some of the circumstances DFID faces in developing countries where respondents may be particularly vulnerable.

Objectives
A team was convened to develop:
1. ethical principles to include in all research and evaluation Terms of Reference
2. a mechanism to support staff commissioning research/evaluation – to help them determine the adequacy of proposed approaches on ethics and share good practice;
3. a system to provide accountability with regard to DFID’s treatment of ethics – recognising that simply providing standards does not ensure they are used.

Methods
We reviewed the documentation of other organisations, looking at ethical principles/standards and support/management systems to uphold these. We also consulted within DFID on the circumstances staff faced on the ground, before producing our own proposals.

Results
DFID developed a set of ethical principles relevant to research and evaluation in developing country contexts and created a system to uphold these. Our ethical principles cover issues of competence, impartiality, conflicts of interest, proportionality, integrity, accountability, enabling participation, confidentiality, cultural sensitivity, dissemination, voluntary participation, vulnerable groups and avoidance of harm.

Application to practice
Issues we faced and decisions we reached concerning the content of our ethical principles will have particular relevance to those working in developing countries. People will also be interested in our experience of upholding these in practice.

Keywords: Ethics; Development; Principles; System; International

Building, with transparency, ethics into all stages of EU Framework Programme ICT Research and Deployment

F. Cunningham
DG INFSO

The ever increasing pervasiveness of Information and Communication Technologies (ICTs) into all parts of our private, working and societal lives, leads to the need to take seriously and to fully address ethical aspects of the research, development and application of ICTs. The European Commission, which holds the remit of the planning, funding, monitoring and evaluation of EU-level ICT research and deployment actions, must take ‘ethical leadership’ and ensure, in a transparent manner, that these actions are ethically sound in terms of their objectives, execution and ultimate application.

This paper will set out the concrete steps taken by DG Information Society to ensure that there is extensive awareness of the range of ethical issues raised by research and deployment of ICTs and to instil ‘ethical thinking’ and behaviour with integrity, by all key stakeholders involved in EU Framework Programme ICT activities. It will also set out how the European Commission will continue to ensure that ethics are at the heart of the EU-level governance of and approach to research and deployment in the fast changing and dynamic domain of ICTs and other areas underpinned by them.
The ethical aspects of working with patient research partners

K. Schipper, C. Nierse, T. Abma
VU University Medical Center, Medical Humanities, Amsterdam, Netherlands

Traditionally patients and their families had two roles in scientific research. They were seen as research objects providing data at the request of researchers or as fundraisers, supporting research and researchers. Nowadays there is an emerging trend in medical and health research to actively involve patients in the research process. Besides qualitative, pragmatic and political considerations, normative arguments lie at the core of this development. Patients are the end-users, and are thus considered stakeholders in research. Their perspectives and experiential knowledge can contribute to the research process and thus lead to research that better relates to patients’ needs and expectations.

The degree of patient participation can vary depending on the amount of control over the research process. Patients acting as objects of research are on the lowest level of participation. Higher levels of participation include patients as advisors and patients as interviewers of fellow patients or moderators of focus groups.

Patients participate in a still higher degree when acting as initiators of research projects or when acting as research partner in a team of professionals. There are a few publications on the collaboration between professionals and patient research partners.

We believe in the value of these collaborations and the value of research partners but some ethical dilemma’s may occur and need to be responded. We did several (evaluation) projects in which we worked together with patient research partners with kidney diseases, muscular diseases, Acquired Brain Injury and intellectual disabilities. Our presentation reflects about the ethical aspects of working with patient research partners like prevention of exploitation, overburdening and the promotion of (relational) empowerment.

Keywords: Patient research partners; Ethical aspects; Exploitation; Overburdening; Empowerment.

Responsibilities in Participation: morality in responsive evaluation

M. Visse – Co-presenters: K. Schipper, C. Nierse, T. Abma (entire session)
VU University Medical Center

Responsive evaluation (RE) contributes to the design and implementation of policy by regarding the evaluation process as a social practice (Abma, 2009; Abma et al, 2009; Abma & Widdershoven, 2008). RE is a democratic process where the evaluator fosters and joins a partnership with stakeholders. The evaluator cares for the inclusion of vulnerable voices and aims to create an ongoing dialogue about differences in perspectives. Therefore, RE promotes specific deliberative ways of dealing with ethical issues as inclusion and participation. This asks for a specific research ethic that goes beyond traditional values of anonymity and confidentiality (Cronqvist, 2003; Giordano, 2007; Taut, 2003; Wiles, 2006).

A responsive, participatory ethic is interpersonal and is strongly connected with the identity, values and relationships of the evaluator and stakeholders. It is based on values as ‘caring about’, integrity, trust, equality, respect, sincerity and transparency. This raises questions about the way this ethic is enacted in the daily practice of RE. Such as: How much should we care? Where do responsibilities begin and end? What to do when values contradict? What competences and responsibilities should be fostered to work with conflicting expectations of stakeholders?

This contribution focuses on how the evaluator develops and expresses a responsive research ethic and what this means for partnerships. Starting point is an expressive collaborative model of ethics (Walker, 2007) and it’s significance for RE. To gain insight in ‘responsibilities in participation’ three narratives will be explored: narratives of identity, relation and values. This is illustrated by a case example of a RE in palliative care. Next, the session provides a practical method that evaluators can use to explore their research ethic: they will reflect on their own ethic and ‘responsibilities in participation’ during an interactive exercise.

Keywords: Responsive evaluation; Morality; Research ethic; Reflexivity; Responsibility
School evaluation

On developing sustainable evaluation culture in schools

L. Nyhus
Lillehammer University College, Faculty of Social Sciences, Lillehammer, Norway

Building capacity for learning and change has become an issue of educational reform and educational leadership in Norway, as well as in other European countries. How to develop a sustainable evaluation culture is one of the challenges school systems are facing. The literature suggests several ‘paths’ schools and leaders might follow for improvement, and building capacity for learning in general is seen as important in the educational sector. In Norwegian politics developing an evaluation- and assessment culture in schools is suggested as vital for learning and improvement and schools have to develop such culture. On the other hand, research has showed that it is difficult for schools to live up to ideals like “learning community” or “learning organization” and also that school innovations rarely sustain. This might mean that these ideal paths or concepts might be difficult to follow or too simplistic by their nature. Maybe development of an evaluation culture should not be seen as a kind of such specific concept or solution. Many researchers now emphasise the need of seeing the complexity in school systems, recognizing the different contexts and the fact that one size does not fit all. In this paper educational systems are seen as constituted of communication, and development of a sustainable evaluation culture is seen in light of this communication perspective, where the point of departure is the ongoing complex responsive processes school actors are facing every day. The paper suggests that a communicative and meta-communicative perspective could be fruitful as a basic competence in the evaluation work.

Keywords: Evaluation culture; Communication

Training for external evaluation in schools

A. Juric Rajh, M. Brejc
National School for Leadership in Education (Sola za ravnatelje), National School for Leadership in Education, Kranj, Slovenia

In 2008 a national project (co-financed by European Structural funds) has been initiated to develop and introduce a system of quality in primary and secondary education in Slovenia. The system will mainly be based on school self-evaluation, but also external evaluation, both combined with a set of quality indicators on a national as well as school level.

While self-evaluation provides an opportunity for the school to reflect on both student outcomes and school development in the light of the goals, targets and improvement strategies as a means of moving the school towards a culture of continuous improvement, external evaluation is defined as a support and encouragement for improving self-evaluation processes. It can contribute to more objective view about the school processes, including setting priority improvement goals, strategies, actions, indicators, leading the process, involving different stakeholders etc. It will also provide information about the quality on a national level.

It is believed that capacity building is of fundamental importance when we are facing the challenges of quality, therefore a training programme for external evaluators was designed. It consists of 7 days of training and intermediate activities during which 32 head teachers and teachers are currently being trained.

The presentation will focus on how core evaluators’ competencies are being developed by using different techniques and approaches, such as team work, document analysis, simulation of practical experience, role-playing and other active methods. Choosing the range of different approaches is even more important since participants from schools with mainly no previous experiences in evaluation are being trained for a pilot external evaluation that will take place at 11 schools in autumn 2010.

Keywords: External evaluation; Training; Evaluators’ competencies
Support of Self-evaluation of Schools in the Czech Republic by the National Project Way to Quality

S. Michek
National Institution of Technical and Vocational Education (NUOv), Department for evaluation tools, Prague, Czech Republic

In line with gradual decentralisation of the education system, Act no. S61/2004 lays down a new obligation for schools to carry out self-evaluation. One of the purposes of self-evaluation is to serve as background information for external evaluation by Czech School Inspectorate. Schools are lacking methodological support from the state and consider self-evaluation to be one of the most problematic issues in the context of developing school educational programmes (which should include self-evaluation). Moreover, there was absence of rules governing the approaches to self-evaluation of schools on the part of schools’ founding entities, inspectors and other stakeholders.

The objective of the project “Way to Quality: Introduction of the System of Self-Evaluation – Self-Evaluation” is to support from May 2009 to April 2012 self-evaluation processes in schools (kindergartens, basic schools, grammar schools, secondary technical and vocational schools, basic art schools, conservatoires and tertiary professional schools), and to define an appropriate combination of self-evaluation and external evaluation in co-operation with schools, their founding entities and the Czech School Inspectorate so as to ensure continuous quality improvement. Project is a national project of the Ministry of Education, Youth and Sports and is financed by the European Social Fund and from the Czech Republic state budget.

In order to achieve this objective, a structured system of proven evaluation instruments and procedures will be developed as part of the project, as well as a support guidance system. The project will also involve support for the sharing of experience among schools, and the development of a system for the training of senior school officials, coordinators and evaluation advisors. Furthermore, a methodology for evaluation the quality of self-evaluation processes and outputs shall be developed, along with an offer for the training of founding entities and the Czech School Inspectorate.

Keywords: Self-evaluation of schools; National project; System support; Evaluation instruments
Interaction between Evaluation Results, Policy and Governance

Evaluation as a step in the policy-making process – The case of EU agencies

J. Toulemonde1, S. Hector1, J. Sylvest1
1Eureval and Lyon University, Partner, Lyon, France
2European Commission, Secretariat-General, Brussels, Belgium
3Ramboll Management, Managing Consultant, Copenhagen, Denmark

At present there are 30 “decentralised” agencies which undertake specific tasks under their own responsibility. They were created over a thirty-year period on a case by case basis. In 2005 the Commission submitted a proposal for improving and harmonising the governance of the “agency system”, but the proposed inter-institutional agreement could not reach the decision stage.

Since then, rapidly growing resources have been devoted to decentralised agencies, and this has raised concerns about the budgetary impact and the relevance of the agency model. Over the same period decentralised agencies have been subjected to a growing number of studies, audits and evaluations. In this context, the Commission, the Council and the Parliament established an inter-institutional working group with a mandate of designing and proposing reforms as far as relevant.

In parallel, the Commission launched a thorough independent evaluation of the agency system, of which the conclusions were meant to feed into the work of the group. This exercise involved extensive consultations with stakeholders, and an innovative reference group was set up in which the three institutions, agencies and scientists could express their views on the evaluation process.

The evaluation was delivered at the start of 2010. Its main focus is the conclusions on the level of the agency system, but there are also findings about individual agencies. Instead of recommendations the report includes options for addressing each main problem identified in the report. Each option is described in terms of time horizon, concerned parties, benefits, risks, difficulties, and the need for further analysis if applicable.

The results of this evaluation are now one of the elements forming the basis for the discussions within the inter-institutional working group. The presentation will elaborate on the involvement of stakeholders in the evaluation process for this show-case of connection between evaluation and decision-making.

Keywords: Policy-making, Stakeholders, Scenarios

Evaluation and the environmental democracy of European cities

A. Bonifazi
Politecnico di Bari, Department of Architecture and Planning, Bari, Italy

Since its early days, environmental assessment aroused high expectations about its ability to insinuate ecological rationalities into worldviews, institutions, and social practices. Besides, basic provisions for information dissemination, public participation and access to justice have been included in most environmental assessment systems. However, the actual contribution of environmental assessment to the democratization of policy making remains highly debated, yet poorly articulated in rather naive (if not entirely implicit) conceptual terms.

This paper aims to broaden the theoretical perspectives on whether and how environmental assessment may foster the democratization of urban governance in Europe. The research relies on cross-fertilization with other evaluation domains, including evaluation in planning and policy and programme evaluation, with a view to expanding epistemological perspectives and to dwelling on the implications of ethics and politics.

Democratization of urban governance is characterized as progress along three complementary processes: broadening environmental citizenship (with due consideration of non-humans and future generations); extending public scrutiny over policies and issues that were previously managed in purely bureaucratic or market-oriented ways; fostering authenticity by problematizing the division of evaluation labour between experts, policy makers and citizens. The alleged influence of environmental assessment is initially discussed by resorting to a typology of evaluations for democracy, which mirror well established root conceptions: the elitist, the participatory, and the deliberative.

Reflections verge on a widespread tendency to keep environmental assessment within the boundaries of highly procedural, merely consultative, and mildly reactive practices. On the other hand, environmental assessment recasts a long-standing friction between claims for self-governance and calls for autocratic handling of environmental issues, against the background of controversial ethical dilemmas. It is suggested that such tensions might trigger innovation in environmental assessment theory and practice, which could prove relevant to other evaluation domains.

Keywords: Environmental assessment; Urban governance; Democratic evaluation; Environmental democracy
S4-09 – Strand 4

Syposium

Sectoral policy evaluation symposium I

Evaluating ETC programmes

M. Kaufmann¹, R. Hummelbrunner², K. Boehme³, M. Novello⁴

¹Evasco s.r.o., Senior evaluator, Prague, Czech Republic
²OEAR GmbH, Senior evaluator, Graz, Austria
³Spatial Foresight GmbH, Senior evaluator, Heisdorf, Luxembourg
⁴Interact Point Wien, Senior evaluator, Wien, Austria

Background

European Territorial Cooperation (ETC) – formerly known under the name of INTERREG – is an important pillar of EU cohesion policy and was introduced as a new Objective for EU Structural Funds in 2007. This Objective is implemented through a range of programmes which are grouped under three strands: cross-border, trans-national and interregional co-operation. While they are subject to the same evaluation requirements as other ERDF programmes, their specific nature poses particular challenges for evaluation (e.g. implementation, impact).

The INTERACT Programme promotes and supports good governance of ETC programmes and provides practical support, trainings and guidance (which includes evaluation). The INTERACT Point Vienna is responsible for delivering services in the south-eastern part of the European Union and was leading several activities in the field of evaluation.

Objectives

The objective of this Symposium is to address core issues for the evaluation of ETC programmes, illustrated by presentations of work recently carried out. And to discuss with presenters and participants about experience gained, lessons learned and challenges posed.

Method

Presentation of three works which have either been commissioned by INTERACT or are considered relevant for the subject.

- “Cross-programme Evaluation” of ETC programmes (Operational and Thematic Aspects) (Böhme K., Hummelbrunner R., Kaufmann M.)
- On-going evaluation of ETC programmes (INTERACT and/or ECORYS,)
- Impacts and beneficial effects of INTERREG programmes and projects (DG REGIO or Panteia)

The presentations will be followed by a joint discussion on all three presentations, which will also address the challenges posed by the evaluation of ETC programmes, in view of their continuation after 2013.

Conclusion/Application to practice

The Symposium will provide an overview of core issues for the evaluation of ETC programmes, information on the results and lessons to be learned from recent relevant work as well as an outlook on future challenges.
Towards more rigorous evaluation of development impact

HIV/AIDS sensitization, social mobilization and peer-mentoring: Evidence from a randomized experiment in Senegal

N. Wagner, J. L. Arcand

The Graduate Institute of International and Development Studies, Economics, Geneva, Switzerland

Using a simple experimental design, Poisson regression techniques and routinely-collected administrative data from health districts in Senegal, we examine whether funding community based organizations can be an effective manner of increasing voluntary testing and counselling and modifying the subsequent behavior of individuals who test HIV-positive. We distinguish between two treatment groups: in a first set of randomly-chosen health districts, community organizations received funding and carried out HIV/AIDS sensitization using traditional social mobilization techniques, whereas in a second randomly-chosen treatment group, they did so by using a new peer-mentoring mechanism; the remaining health districts were assigned to the control group and received no funding. Our results indicate that: (i) funded peer mentoring doubles the number of individuals who get tested, who follow pre-test counselling and who pick up their test results, whereas funded traditional social mobilization appears to be ineffective; (ii) neither traditional social mobilization nor peer-mentoring increases the number of individuals who tested positive and who picked up their results; (iii) both traditional social mobilization and peer mentoring increase the number of partners of HIV-positive individuals who got tested, with the effectiveness of traditional social mobilization being much larger than that of peer-mentoring for men and marginally so for women; (iv) both peer mentoring and traditional social mobilization significantly increase the number of HIV-positive individuals who follow post-test counselling, with peer-mentoring increasing post-test counselling among men and traditional social mobilization being significantly more effective at increasing post-test counselling among women.

Keywords: HIV/AIDS sensitization; Social mobilization; Peer-mentoring; Randomized control trial; Poisson regression

On measuring and bridging through impact evaluations: lessons from Agence française de Développement’s experience

J. Delarue, T. Bernard, J. D. Naudet

Agence Française de Développement, Evaluation Unit, Paris, France

The number of impact evaluations related to development interventions has tremendously increased over the recent years. Central in this evolution is the widely shared idea that impact evaluations constitute a unique opportunity to bridge the gap between researchers’ and policy makers’ communities, and contribute to development effectiveness through better knowledge. Thus, the most famous example in Mexico showed how researchers were able to quantify the effect of Progresa on school enrollment and health outcomes, which has led to important policy consequences in Mexico itself, but also in other Latin-American countries where the approach has been replicated. Yet, while several Progresa-like success stories exist, not all impact evaluations are as successful in Bridging researchers and policy-makers, and in measuring whether a program has reached its objectives. In fact, interventions are more or less suited to actual measures of their effectiveness even among the ones with clearly defined treatment and control groups. Further, bridging often requires extra efforts to align the often diverging interests of researchers and policy makers.

The discussion relies on two on-going studies, to be completed by the end of 2010, assessing (i) the impact of micro-credit on rural poverty in Morocco, and (ii) the impact of health insurance on morbidity in Cambodia. Both rely on experimental designs and on extensive data collection involving more than five thousand households. Each study is conducted by high-level researchers from northern universities, associated with experienced data-collection firms.

Yet, both studies reveal important difficulties in measuring impacts, given low compliance rates and their limited time-span. Overall, expected results will probably be limited to identifying households’ behavior when a given constraint is relaxed. These contributions are of crucial interest for the comprehension of development issues, yet, the programs’ implementing agencies may feel frustrated by the produced results, and bridging may consequently be limited.

Keywords: Impact; Method; Policies
Challenges of Impact Evaluations in the context of German Technical Cooperation

S. Krapp
German Technical Cooperation (GTZ), Evaluation, Eschborn, Germany

Projects in development cooperation (DC) deal with complex reforms and change processes whose effectiveness is often difficult to visualise. The actual success of DC is often discussed. The public, politicians and partner countries all expect reliable statements on the effectiveness, sustainability and efficiency of DC. Not only when budgets are tight, DC is required to provide credible proof that it uses available funds effectively, efficiently and thereby making a relevant contribution to the economic and social development of partner countries. The first line of questioning concerns the impact of projects and the extent to which successful development in a given country is directly or indirectly related to the supported projects and programmes. At the latest since the formulation of the Millennium Development Goals and the Paris Declaration, the discussion of the quality and effectiveness of DC has gained new importance. The development of standardised procedures and instruments for measuring and attributing the results of DC play an increasingly important role in this context.

Against this background GTZ (German Technical Cooperation) aims to meet the challenges to provide evidence of impact of its development measures. First, by implementing structural prerequisites of its evaluation system: establishment of an independent evaluation unit clearly separated from the operating departments, application of different evaluation instruments (self-evaluations, decentralised evaluations, independent evaluations, joint evaluations, external evaluations). And second, by applying respective rigorous evaluation designs specifically in its final and ex-post, but also ex-ante evaluations carried out every year with a specific sectoral focus. The goal is to accomplish as many as possible high quality impact evaluations as rigorous, standardised and efficiently as possible. The hereby gained experiences provide conclusions for further application of methodologically ambitious evaluations in practice, which can be further discussed with the attendees of the presentation.

Keywords: Rigorous Impact Evaluations; Development Cooperation
Evaluation use, evaluation impact III

Knowledge transfer partnerships and community development: using participatory approaches to enhance the impact of evaluations in health and education in England

S. Vougioukalou
c/o Bridget Carroll Canterbury Christ Church University Department of Social Wo, Canterbury-Kent, United Kingdom

The UK government is currently encouraging universities to increase the social and economic impact of their research by forming knowledge transfer partnerships with community organisations and local businesses. This paper discusses the current impact of knowledge transfer partnerships between the newly formed Service Evaluation and Development Group at the Department of Mental Health, Community and Mental Health at Canterbury Christ Church University and organisations in the education and health & social care sectors.

The evaluation of seven organisations were undertaken between 2007 and 2010. These include statutory and voluntary organisations that provide services to pupils on at the brink of school exclusions, children with special needs and their families, people with dementia, cancer patients and other vulnerable groups. The methodology used in the evaluation process included a wide range of quantitative and qualitative tools such as action research, participatory decision-making, appreciative inquiry, participant observation, focus groups, interviews and questionnaires.

Increased participation from service-users and personnel including administrative support staff, specialised staff and managers increased not only the transparency, efficiency and timeliness of the evaluation but also the sustainability of the organisations. Furthermore, it increased the clarity by which relevant evidence was brought to the attention of decision makers, it opened up debate within the organisation and amplified the voices of neglected groups, thereby contributing to a transparent and sustainable model for community development. Methodological restraints and possible ways to overcome these with the use of new technologies and online collaborative platforms are also discussed.

Program evaluation evidences: utility or futility? An empirical analysis from Canadian public servants

M. Lamari¹, J. Jacob¹, M. Ouimet¹, J. Turgeon¹

¹Enap, Créee, Québec, Canada
²Laval University, Political science, Québec, Canada
³Enap, Gépps, Québec, Canada

Our paper deals with the knowledge utilisation produced by program evaluation within Canadian public services. Using data from a survey conducted, in 2009, among a representative sample of 1,600 professionals, working for Quebec government, our research empirically answers three complementary questions:

i) Among all the knowledge produced and consulted in order to improve governmental decision making, what is the contribution of evidences produced by program evaluation?

ii) What are the determinants of the utilisation of evaluation results, and how do they differ from other determinants related to other knowledge utilisation sources?

iii) To what extent does the utilisation of evidence produced by program evaluation brings changes and innovation on public policy?

To answer these questions, advanced statistical analyses were carried out (multinomial logistic regression, OLS regression, correlation and ANOVA). Our hypotheses are inspired by two theories: i) knowledge utilisation approach (Weiss, 1979; Rich, 1997, Landry et al., 2003) and ii) the advocacy coalitions approach (Sabatier and Jenkins-Smith, 1993).

Our findings suggest that only 60% of civil servants are interested on the use of program evaluation results. This interest is unequally distributed among departments and ministries. Public servants operating in finance, justice and economic development departments are less committed to consider program evaluation results in their decision-making process. Public servants operating within health, education and agriculture departments show bigger interest toward them. Most important, the likelihood of the use of program evaluation is more important when professionals have a post-graduate degree, are trained in social science/humanities, and are familiar with advanced research methodologies. Age, gender and professional experience do not show significant statistical coefficients. Our analyses derive new avenues for public management and identify some implications for future research on the utilisation of evidence produced by program evaluation.

Keywords: Evaluation use; Evaluation utilization; Knowledge utilization; Advocacy coalitions; Empirical research
Gender and evaluation

Gender and evaluation: building bridges of theory and methods

K. Hay, R. Govinda, V. Ramachandran, K. Renu, V. Magar

1IDRC, Evaluation Unit, Delhi, India
2Ambedkar University, Development Studies, Delhi, India
3Educational Resource Unit, Director, Delhi, India
4Society for Health Alternatives, Founding Member, Baroda, India
5reach, Director, Delhi, India

The practice of evaluating women’s empowerment programmes or the ‘gender’ component in development has expanded in recent years in South Asia, making it important to document and theorize on gender, development and evaluation. This symposium critically examines the theories influencing, and the methods being used to evaluate gender and development in South Asia. Drawing on several cases, the symposium identifies methodological challenges in evaluating women’s rights oriented programs. The paper by Radhika Govinda maps the theoretical terrain of measuring change in gender relations in South Asia and explores how it compares and contrasts with similar efforts in the North. The paper by Vimala Ramachandran, examines gender and social equity in elementary education in India. The paper draws on efforts to assess the effectiveness of educational programmes on reducing gender and social equity gaps in participation. The paper critically reflects on the methodologies adopted and how evaluations were received and used in the system. The paper by Veronica Magar, examines the challenges of evaluating women’s empowerment programs in the context of conflicting ideologies and conceptual frameworks on gender and sexuality. It examines two evaluations using case mapping approaches on interventions with sex workers in India and rural women in Bangladesh and a participatory evaluation with a women’s rights organization in India and critically reflects on methodologies. The paper by Renu Khanna, draws upon her experience evaluating reproductive and child health interventions in India. The paper examines how the evaluations looked at gender, rights and entitlements and discusses methodological choices in capturing issues of gender and power and in promoting gender equity within the programmes being evaluated. The Panel Chair, Katherine Hay, argues that the field of evaluation has inadequately drawn from feminist research, both theoretically and in practice and attempts to identify theoretical and practical opportunities to strengthen evaluation theory and practice.

Keywords: Gender; Women’s rights; Methodologies; South Asia
Politics of evaluating in and on Violently divided societies

Perspectives on the politics of evaluating research in and on violently divided societies

C. Duggan, K. Kelly, K. Bush

1 International Development Research Centre, Evaluation Unit, Ottawa Ontario, Canada
2 Institute for Social and Economic Research Rhodes University, Centre for AIDS Development Research and Evaluation, Grahamstown, South Africa
3 International Conflict Research Institute, University of Ulster, Londonderry, N. Ireland / U. K.

Evaluating the reach and effects of applied research for social development in violently divided societies is a complex enterprise, fraught with particular challenges: some of these are shared or common to any social change intervention involving research, including non-linear paths towards change, attribution of results and time-lags between knowledge production, uptake and mobilization. In violently divided societies, these challenges can be compounded or exacerbated by the contextual particularities that characterize these settings: security of evaluators and evaluation subjects, political manipulation of evaluation findings, restrictions around using good participatory process and limited access to necessary information and institutions, to name but a few. Most conventional approaches for evaluating research programming in violently divided societies are ill-suited for understanding and assessing social change. There is a need for new approaches that are structured and calibrated for the particular, and often extreme, conditions within violently divided societies. This panel will explore how program evaluation may achieve the dual goals of evaluation for evidence and evaluation for process benefits. Panelists will draw on findings emerging from a joint International Conflict Research (INCORE)/International Development Research Centre (IDRC) exploratory research project. Three contributions to the panel are contemplated: A paper examining the politics of doing research and program evaluation in these contexts and challenges for research uptake into evidence-based policymaking processes; A paper exploring the political implications of methodological orthodoxy and their effects on the relationships underpinning the evaluation process; and a paper assessing the institutional politics around evaluating research in a specific sector (HIV-AIDS research) in violently divided societies.

Keywords: Complexity; Conflict; Research; Violence
Ethics and independence

Project Ethics
B. ta Larsson 1, L. Barrett-Smith 2
1 BIM Larsson & Associates, Evaluation Firm, Edmonton AB, Canada
2 Alberta Innovates – Health Solutions, Research Ethics Initiatives, Edmonton AB, Canada

Introduction
Evaluations can play a major role in building inclusive communities and supporting democratic processes. But how do we make sure that everyone’s voice is heard while still respecting and protecting individuals involved in evaluation projects? What knowledge and skills do evaluators need to ensure ethical conduct in their project work?

The issue
Growing in number and complexity, evaluations often deal with at-risk or vulnerable populations in intrusive ways. Historically, the participants in evaluation projects have not been given the same protection through ethical scrutiny as research subjects, nor have there been guidelines or processes in place to assist addressing ethical risks to participants. Often, evaluators and commissioners of evaluations have to rely on evaluators’ professional ethical guidelines and codes of conduct to ensure that project ethics are considered. These types of guidelines are neither comprehensive nor specific enough in terms of practical application of ethical concepts.

Approach
Project ethics involves the integration of ethics considerations from the planning through to the reporting stages of a project to ensure that the rights of the people involved are respected. During the roundtable discussion project ethics concepts and two practical ethics decision-support guides will be introduced to assist the development of essential skills for evaluation practitioners. Discussion will focus on the application of a framework to enhance consistency in identifying and managing ethical risks to participants in projects. Two guides support this framework: 1) Six Ethics Considerations for use in developing evaluation projects, and 2) an Ethics Screening Tool to determine appropriate review requirements and identify ethical risks for project participants.

Objectives
Attendees will discuss: Why ethics matters in evaluation? What is project ethics? How to identify and address ethical issues in evaluation projects using a practical framework with guides.

Keywords: Project ethics; Ethical risk; Tools

Institutionalizing Evaluation in South Asian Perspective: A Case Study of Pakistan
G. Mustafa
Earthquake Reconstruction and Rehabilitation Authority, Project Evaluation Cell, Islamabad, Pakistan

This paper envisages taking an overview of the present status of evaluation in Pakistan. Planning Commission of Pakistan is responsible for (a) preparation of national plans at periodic intervals for the economic and social development of the country; (b) making assessments from time to time of the human and material resources of the country; and (c) to prepare the Annual Development Programme (ADP) within the framework of the national plan and on determination of priorities. It has been assigned the role of monitoring and evaluation of policies, programmes and projects, which is carried out by one of its wings known as Projects Wing. The Projects Wing has recently been strengthened by provision of additional staff, equipment, etc. It works under the leadership of one of the members of Planning Commission, who directly reports to the Deputy Chairman, who chairs the Central Development Working Party (CDWP), which is empowered to approve projects costing up to Rs. 1,000 million. The Chairman of Planning Commission is the Prime Minister of Pakistan. The Prime Minister also chairs the Executive Committee of the National Economic Council (ECNEC), the highest forum for approval of projects in the country.

The Projects Wing of the Planning Commission of Pakistan is striving to introduce a culture of results based management (RBM) in the country. It issued a document in 2008 entitled “Guidelines for Project Management”, which emphasized the importance of RBM for the first time. The Planning Commission also runs Pakistan Planning and Management Institute (PPMI), which endeavors to impart training to officers of federal ministries and provincial governments in project planning and implementation.

As a result of the steps taken by the Planning Commission, some improvement has been seen in the implementation of projects. However, there is much still to be desired.

Keywords: Evaluation; South Asia; Pakistan; Planning Commission
Empowerment

Evaluation of Capacity Development Interventions
G. Reyes Suarez
UNICEF, Social policy monitoring and evaluation, Cairo, Egypt

One of the major strategic inputs in terms of programme assistance in Egypt is focused on capacity building activities, particularly training, in order to advance institutional development, technical empowerment information management and improve quality of services. UNICEF Egypt Country Office (ECO) Programme of Cooperation has been supporting training and capacity enhancement activities expected to result in improved skills, practices, awareness raising, and behaviors changes.

For this reason ECO is working to institutionalize the evaluation of its capacity development activities. This step is an attempt to ensure systematic accountability on learning and capacity development of counterparts. As a result of using different evaluative approaches, evaluation tools have been developed to track the absorption of knowledge from the participant and the impact on her/his works after attending a training. The technical approach and methodology used is based on Kirkpatrick framework adding some other utensils such as Michael Scriven training evaluation checklist.

As part of this initiative, the Social Policy, Monitoring and Evaluation section (SPME) is also developing a capacity building matrix that will also provide technical support to the other ECO sections in order to translate capacity building objectives into monitorable indicators that can be assessed using various frameworks and other tools. In this regard, this section is developing a model framework for such capacity evaluation in our future interventions.

Keywords: Capacity development; Accountability of resources; Generate evidence on the achievement of results and institutional performance

Evaluation, civil society empowerment and accountability Lessons from Burkina Faso
D. Thieba
GREFCO, Burkina Faso

In Burkina Faso like in most countries in transition, policy matters are formulated without a significant involvement of civil society. During the last decade the State and funding institutions have insisted on the participation of civil society in mid term evaluations and periodic reviews of SFPR (Strategic Framework for Poverty Reduction) formulation. In 2005, that is five years after the formulation of such requests, the situation had remained the same. Participation in dialogue and reviews remained a wishful thinking because the civil society had no say in the adoption of reforms and were not even considered.

The use of a monitoring and evaluation tool and the need for skilled people for relevant evaluations have contributed to significantly strengthen civil society in the last five years. An independent expertise has been made available to civil society to assist them in the monitoring and evaluation of policies and programs and putting emphasis on relevant indicators for impact assessment, efficiency, and analysis of results. Admittedly, this approach has helped civil society to strengthen its abilities in discussions and supporting proposals, which reduces the imbalance in the relationship with the State in public talks. Civil society has gained some respectability. Changes have even been noted in policies in some areas (improvement of indicators to measure poverty, adoption of new measure).

This article has a double objective:
(i) To describe how the monitoring and evaluation tool can be useful to the actors of civil society in the talks on policy matters.
(ii) To explain why this approach has contributed to improve the respectability of civil society in decision taking even if there are still some limitations.

Keywords: Evaluation; Civil society; Policy dialogue; Empowerment
Evaluation and the care sector

Participative Efforts in Evaluation – Voices of the Less Fortunate

T. Riisom, M. Lindgaard, S. Bohm

1 Deloitte Business Consulting AS, Copenhagen, Denmark
2 Ministry of Finance, Copenhagen, Denmark

A Danish initiative is testing the possibilities of creating comparable information about the satisfaction among elderly and disabled citizens. The experience from the pilot projects conducted in 2009–2010 is based on 2,300 personal interviews, 5,700 postal interviews and 1,800 telephone interviews among the users of home care, care centres and residential accommodation.

The presentation is tripartite: In the first place, the background of the initiative and its connection to the current quality improvement reform are accounted for. Furthermore, the democratic and management motives of testing methods that increase the target groups’ access to express their views on the public service they receive are accounted for.

Secondly, the presentation accounts for the target groups’ special characteristics and the importance of these characteristics to the citizens’ access to exercise basic democratic rights and, specifically in this context, the access to express their views on and assess the service provided.

Finally and thirdly, the presentation accounts for the methodological challenges necessary to overcome during the pilot projects. The methodical part of the presentation will summarise the experience gained from the pilot projects with a special focus on the design of question frames, the use of varying ways of interviewing and the determined reasons for defection of users in the target groups.

Keywords: Welfare sector; User satisfaction; Methodology; Pilot project

Outcomes of physical autonomy in Post Acute and Long Term Care

J. Abreu Nogueira, A. Giraio, I. Guerreiro

Ministry of Health, Mission Unit for Continuous Integrated Care, Lisbon, Portugal

Background

The use of an integrated tool for bio-psychosocial evaluation of users in need of post acute and long term care and their outcomes registered on a paper-free on-line web based system of data management allows the continuous process of results’ assessment. It allows real time results at a national, regional, local, and unit level, making benchmarking possible to obtain with a large number of registries.

Objectives

To characterize the degree of physical autonomy on admission and discharge of the users of the National Network for Continuous Integrated Care (RNCCI) as well as characterize their age group during 2009.

Methods

Analyze the reports on the on-line web based system, related to age groups of users and their degree of autonomy. The bio-psychosocial evaluation tool, related to physical autonomy has four degrees related to scores obtained: Incapacity, dependent, autonomous and independent.

Results

Pertaining to physical autonomy, the data management system allowed to identify that in 2009, 80.5% of the referrals to the RNCCI, concern users aged 65 or more and 42% are aged 80 or more. Evaluating 5,784 registries for admission and the same number on discharge, it was possible to obtain 27% decrease of incapacity, 16% decrease in dependents, an increase of 312% of autonomous and 393% increase of independents. These results are obtained in a population of users that on admission 43% have incapacity and 52% were dependents, representing 94% of the users.

Conclusion/Application to practice

This ongoing monitoring allows the implementation of a continuous improvement policy and allows carers to be engaged on performance management and benchmarking between them. This data management system allows obtaining a large number of registries for analysis otherwise only possible in prospective or retrospective studies specifically designed for this purpose.

Keywords: Outcomes; Long-term care; Evaluation; Disability
Post Acute and Long Term Care: instrument for evaluating outcomes
J. Abreu Nogueira, M. J. Girão, I. Guerreiro
Ministry of Health, Mission Unit for Continuous Integrated Care, Lisbon, Portugal

Background
Data from post-acute and long-term care are not widespread in the literature and are essential for international benchmarking. There are different instruments to characterize the needs for those citizens eligible to promotion of autonomy due to temporary or permanent disability. International recommendations point to an integrated bio-psychosocial instrument for initial and ongoing evaluation to assess results of care and research.

Objectives
To develop an integrated instrument for bio-psychosocial evaluation and its implementation on the evaluation of users in need of post acute and long term care and outcomes of care.

Methods
For the creation of a single tool it was established a partnership with a university to assemble different instruments of evaluation. The single tool should include different parameters regarding physical autonomy, activities of daily living, mental and cognitive status, nutrition, falls and social. At the end the instrument gives a score for each parameter analyzed and a combined score. There was a focus on professional training for the new tool.

Results
After the necessary adaptations of the instrument with the university it was implemented in the National Network for Continuous Integrated Care (RNCCI), to be used for evaluating users in need of post acute and long term care and their outcomes. The introduction of the tool allowed ongoing monitoring of different parameters in a holistic approach gathering bio-psychosocial data from a single tool and benchmarking. Evaluation of the users with the tool is made on admission, monthly for post acute care, every three months for long term care and before discharge.

Conclusion/Application to practice
Evaluating users of post acute and long term care based on a holistic approach with a single bio-psychosocial tool allows multidimensional analysis pertinent for research regarding care needs and outcomes.

Keywords: Long-term care; Tools; Evaluation; Bio-psychosocial
Evaluation Systems in the Public Interest?

L. Goodyear, Ph. D, 1, P. Dahler-Larsen, Prof. 2

1 Program Director, National Science Foundation, Division of Research on Learning in Formal and Informal Settings, Arlington, USA
2 University of Southern Denmark, Political Science and Public Management, Odense, Denmark

A striking development in the field of evaluation in recent years is the emergence of evaluation systems (Leeuw and Furubo 2008); mandatory, comprehensive, repetitive organizational mechanisms for the production of evaluative information (Rist and Stam 2006). Evaluation systems include systems of performance management; audit, inspection and oversight; accreditation; and monitoring.

Although advocates of evaluation systems argue that they can enhance systematic approaches to both accountability and organizational learning, literature suggests that evaluation systems may have a number of negative or unintended effects, including that they enhance single-loop learning but hinder double-loop learning; provide only procedural assurance; focus on performance but not on the assumptions undergirding existing policies (Leeuw and Furubo 2008: 165); incur large hidden costs (Power 2005: 335); are marred by a performance paradox so that more measurement does not lead to more quality (van Thiel and Leeuw 2002).

Relatively little is known about how evaluation systems function within social and political contexts. In this session, one European and one American will interview each other and then discuss how evaluation systems work in their respective contexts. The interviews will start with these three questions:

1. From your position/viewpoint and in your context, what do you see as the most important driving forces behind the emergence of evaluation systems?
2. Is there, as you see, a need for a new and better understanding of the side effects of such systems, including, say, their impact on the daily life of employees, their constitutive effects, and their socio-political impact?
3. What kind of public debate do you see around you about evaluation systems, and is any aspect of the debate lacking if our standard is that evaluation systems should be in the public interest?

Keywords: Evaluation systems; USA; Europe
Evaluation competencies

Evaluation competencies: Definition, Teaching, Practice

J. Toulemonde¹, S. Kihm², C. Tourmen³, J. U. Hense⁴

¹ Lyon University and Eureval, Associate Professor, Lyon, France
² Saarbruecken University, Coordinator of the Master Evaluation, Saarbruecken, Germany
³ ENESAD (French National Institute of Higher Education in Agronomy), Assistant Professor, Dijon, France
⁴ LMU München Department Psychology, Empirische Pädagogik und Pädagogische Psychologie, München, Germany

This session contributes to building bridges between the world of education and training and that of evaluation practice. It is designed by the European Collaboration on University Study Programmes in Evaluation (USPE), a network of six European university programmes delivering master degrees in evaluation and deeply concerned with competencies. USPE members co-operate since 2004. They monitor international efforts towards defining (essential) evaluation competencies. They support the initiatives of evaluation societies in this area, both at national level and European level.

The first presentation by Stefanie Kihm, Saarbrücken University (Germany), reports on the related activities and achievements of the USPE Collaboration during the last two years. The works started by building and passing a comprehensive internal questionnaire covering about 100 competency items (based on international works such as Russ-Eft et al., 2008). The second step is a simplified questionnaire focusing on essential competencies, which is meant to describe the competency profiles of education and training programmes in a comparable manner across Europe. The draft tool will be presented and discussed in Prague.

The second presentation by Jan Hense, LMU München, reports on the on-going DeGEval works on teaching and training evaluation competencies. Based on the ‘Recommendations on Education and Training in Evaluation – Requirement Profiles for Evaluators’, published by DeGEval in 2004, the current project aims at developing accompanying guidelines and recommendations for the didactics of evaluation.

The third presentation by Claire Tourmen, ENESAD Dijon, elaborates upon the concept of ‘competency in situation’, following a four level hierarchy of capabilities: (1) doing something, (2) doing something in an efficient, rapid, reliable, and elegant manner, (3) adapting one’s way of doing something in order to fit in with specific situations, and (4) finding a way of doing something in a completely new situation.

Overall, the presentations and discussions are meant to promote transparency and dialogue about evaluation competencies on both supply and demand sides.

Keywords: Competencies
The Interface of policy and research

Proving the impact of R&D support on competitiveness: the case of the SME specific measures under the EC Framework programmes

G. Steurs, V. Duchene
IDEA Consult, Competitiveness and Innovation, Brussels, Belgium

Proving the impact of R&D-support on the competitiveness of those supported, is a very difficult exercise. The nature of the EC Framework Programmes for R&D, based on pre-competitive collaborative research projects means that there are many steps, and complementary inputs downstream of the research, before commercial or other types of return can be realised. There are the problems of attribution as well as with the many ‘spillovers’ which flow from a research action. Data availability, especially tracking research over the longer term is also a severely limiting factor.

This paper presents the methodologies that have been used to make an impact assessment of the SME-specific measures of the Fifth and Sixth Framework Programmes for Research. The rationale of the SME-specific measures is to involve SMEs that do not have the capacity to carry out the required research on their own. The underlying ‘philosophy’ is to provide funding to SMEs to allow them to outsource research to RTD organisations, which would in turn provide R&D services for the SMEs.

In order to assess this impact, different and complementary methodologies were applied. Online surveys were set up that addressed to both participants and applicants whose proposal passed the evaluation threshold but was not selected for EC-funding. 45 case studies provided case-based, empirical evidence on the impacts of participation on SMEs. There were additional interviews and a workshop with stakeholders. Finally, the AMADEUS database was used with economic and financial data on millions of EU companies, to evaluate the impact of the SME specific measures on the participating SMEs.

Each of the above mentioned methodologies has its strengths and weaknesses. However, the approach we followed by combining the evidence of the four evaluation tools provided an evidence base that is much stronger when compared to an approach based on one single methodology.

Keywords: Evaluation methodologies; R&D-support; Impact assessment; SME; Competitiveness

At the cross-roads of national policy objectives and the global operational environment

K. Lahteenmaki – Smith, M. L. Niinikoski, P. Kutinlahti
1 Net Effect Ltd., Policy advice group, Helsinki, Finland
2 Net Effect Ltd., Mikonkatu 15A, Helsinki, Finland
3 Ministry of Employment and the Economy, Innovation Department, Helsinki, Finland

Our paper presents some of the findings of the institutional evaluation of VTT Technical Research Centre of Finland. The evaluation has been commissioned by the Ministry of Employment and the Economy (MEE). We have identified measures and recommendations for reforms that support the effective renewal of VTT’s expertise, as well as developing VTT’s management by results, performance and management practice. In order to achieve these goals, we have designed a plan centred on the three key dimensions of VTT’s role and the analysis thereof, namely, VTT’s role as part of the innovation system; VTT as part of the MEE Group and VTT’s operational role. The aim has been to produce information and analysis that will support practical decision making with respect to each of the dimensions of VTT’s innovation environment, i.e. the international and global level, the national level and the local/regional level. The paper will elaborate some of the methodological challenges in evaluating innovation policy, where the steering and guidance of the ministry, the operational and strategic autonomy of a research organisation and the needs and expectations stemming from the innovation environment meet.

The perspective of the public sector is institutional in nature and could even be viewed as shedding light on the ‘Finnish national interest’, while the perspective provided by stakeholders representing the research field and enterprises necessarily has a stronger link to thematic bodies of expertise and to a client, partner or competitor perspective. We also report some of the findings on the tension faced by research organisations such as VTT from the perspective of external elements of the innovation system. The ways in which these tensions can be assessed in an evaluation is one the key issues discussed in our paper.

Keywords: Innovation policy; Public Research Organisations
Evaluating Research and Development Regional Policies

G. Garau, F. Mazzeo Rinaldi, A. Spano, G. Mandras

1 University of Sassari, Sardinia Regional Evaluation Unit, Sassari, Italy
2 University of Catania, Sardinia Regional Evaluation Unit, Catania, Italy
3 University of Cagliari, Sardinia Regional Evaluation Unit, Cagliari, Italy
4 University of Sassari, Department of Economics, Sassari, Italy

It is widely recognized that research and development (R&D) contribute directly to our level of prosperity and the well-being of individuals and society, increasing quality of life in general. European Institutions are working in order to establish the EU as a leading knowledge-based economy. In this context, Sardinia Region has invested a considerable amount of public resources on R&D over the past years, even above national average between 2000–2005. Despite this financial effort, almost nothing has been done to evaluate the effects of this policies.

The main aim of this paper is to evaluate the effects of R&D policies implemented in Sardegna over the past 15 years, according to the characteristics of such intangible investment: long-term performance and interaction with other factors of production. The idea is also to assess the investment dimension related to the R&D policies targets.

In order to assess the economic impact and estimate the relationship between R&D investment and productivity, the methodology here proposed is based on a computable general equilibrium model (CGE). The logic underlies the method is to understand whether and when an exogenous shock leading to an initial perturbations of the general equilibrium state, can be reabsorbed, bringing back the system to its long term equilibrium path. The model explain how increase in productivity will be distributed among economic sectors, showing the entire path of economic effects activation, as a result of the public investment.

The analysis of the distribution of R&D public resources among economic sectors and the development and the implementation of the model, will provide invaluable information to better understand whether and where investments have led to the structural changes expected and whether they are consistent with the decision maker’s objectives.

Keywords: R&D policy; Impact Evaluation; Computable General Equilibrium Model
S2-05 – Strand 2 Paper Session

The role and use of evaluations in the policy making process

Ten reasons not to monitor policy implementation, and what it means to evaluations

L. Hermans
Delft University of Technology, Faculty of Technology Policy and Management, Delft, Netherlands

The role and use of evaluations in political decision-making has been subject to much study in the past. This has shed light on the various ways in which evaluations are (mis)used by decision makers, and reasons for these uses. Past research has not yet addressed in sufficient depth the aspects of policy implementation and monitoring, as logical analytical precursors to (ex-post) policy evaluations.

Exploring the link between evaluation, politics and implementation and monitoring yields useful additional insights for the debate on the role and use of evaluations. Monitoring data are being collected through different systems, contributing for instance to administrative statistics and bureaucratic accounting schemes. These data typically address policy inputs, such as budgets or staffing, and states of a policy system. What happens in between policy inputs and changes in system states remains a black box. Looking at monitoring through a political science lens provides numerous reasons for not monitoring actual policy implementation. These reasons will be addressed in the presentation, ranging from political disinterest, to implementation problems and issues of “trust versus control” in policy administrators and professionals.

If policy implementation is not monitored, the analytical quality of ex-post evaluations is impaired. Evaluation studies still can be useful, but policy evaluations based on complete and tailor-made monitoring data are unlikely. One implication is that policy evaluations cannot claim to have a full understanding of policy success or failure. Another implication is that evaluators cannot blame others – notably the administrators and bureaucrats who fail to supply the data – for not being able to do so.

Keywords: Monitoring; Implementation; Politics; Evaluation utilization

The contribution of evaluations to a structural public sector reform in Denmark

O. Rieper, L. Olsen
Akt Danish Institute of Governmental research, Akt, Copenhagen, Denmark

In Denmark a large public sector reform was implemented starting 2007. Among other elements, the reform reduces the number of local governments from 272 to 98 in number, and transferred the delivery of many services from the regions to the local governments (the municipalities). One of the political aims of the reform was to obtain better coordinated and wholeness-oriented services by placing the service delivery under one political authority, namely the local governments (in Denmark the local governments are a powerful tier of government with own elected council and tax income and the main responsibility for welfare services to the citizens).

The paper addresses two main questions. 1. What role – if any – evaluations and research played in the preparation phase of the reform. Evaluation results might help clarifying questions and concerns among policy actors during the preparation phase, or the results might be purely symbolic by only serving as legitimation for the decision, or evaluation results might serve as political ammunition in negotiations. 2. What were the intended uses of the research and evaluation projects that were launched around 2007 on the implementation of the reform? And what has been the actual uses – if any – by spring 2010?

The two questions are to be answered on the basis of results from existing research, information from key actors and a recently finished large evaluation on specialized social institutions.

The paper tells a story on what roles evaluations and research play in a transition process of a huge public sector reform at the level of local governments.

Keywords: Structural reform; Public sector; Turbulence; Change; Evaluation policy change; Structural reform; Change; Turbulence
Evaluation among governments through Peer Review: reducing external pressures and increasing rationality and appropriation in public policy

Design

A. García Femenía
Instituto de Desarrollo y Cooperación, Universidad Complutense, Madrid, Spain

The following presentation is based on a research conducted by the University Institute for Cooperation for Development (Complutense University of Madrid) and the International Foundation for Public Policies and Public Administration, and it was developed in the context of the VI Euro-LatinAmerican Heads of State and Government Summit that will be held in Madrid on 18th May 2010.

The objective of this research was to analyze the possibilities of establishing a mechanism of peer review among LatinAmerican and European countries as part of the evaluation of social policies, particularly public policies for social cohesion.

The three main international peer reviews mechanisms have been analyzed: peer review in Organization for Cooperation and Economic Development, peer learning in the European Open Methods of Coordination, and African Peer Review Mechanism. Several evaluations, particularly in the case of the Open Methods of Coordination in different EU policies have been taken in consideration in the research.

The main results of the study can be summarized as following: Peer review, or peer evaluation among countries, at a governmental level, and with a participatory methodology, can contribute to enhance the learning capacity of public servants and policy makers; and reduces the negative influences of other types of diffusion of public policies, as external pressure or international legitimacy. It introduces rationality in the way public policies are selected and designed by the policy makers, therefore also reducing other influences in the learning procedures like the cognitive heuristics.

Evaluation among governments as an instrument to design public policies can therefore be a positive mechanism to ensure appropriation and should be supported by the international cooperation, particularly in medium income countries, as the LatinAmerican countries.

Keywords: Rationality; Appropriation; International cooperation; Governments and policy makers; Peer review
Evaluation in childrens and youth services

Katimavik; Canada’s National Youth Volunteer Service Program: Using Evaluation to Improve Cost Effectiveness

H. Cummings1, P. Nadine2, M. J. Deblois1, S. Healey3, A. Cummings3, I. Brisebois2
1University of Guelph, Environmental Design and Rural Development, Guelph, Canada
2Katimavik, Program Evaluation, Montreal, Canada
3Harry Cummings and Associates, Program Evaluation, Guelph, Canada

Background
In Canada, the Katimavik program is a national volunteer service program for youth between the ages of 17–21. Its mission is: “Engage youth in volunteer service and foster sustainable communities through challenging national youth service programs”. In recent years it has placed in excess of 900 youth in communities across the country. The program is funded nationally (Heritage Canada) and has as its objectives: “to contribute substantially to the personal, social and professional development of the volunteers; to promote community service; and, to offer a diverse experience fostering a better understanding of the Canadian reality.” (www.katimavik.org)

The program has been in existence for 33 years, has received very high approval rating from participants and has impacted their knowledge, attitude and practices significantly (Harry Cummings and Associates, 2010). It has been evaluated annually using a pre–post questionnaire format administered to all participants, supplemented by regular focus groups.

Objectives
Continuous pressure on budgets and the desire to innovate led Katimavik to explore other forms of delivery. The objective of this evaluation was to determine if the new 6 month program could have impacts similar to the classic 9 month program.

Methods
The methods used included a comparison group of non participants, pre and post questionnaires for participants, focus groups and focus group questionnaires and interviews. The evaluation included reliability and validity testing for questionnaires, multivariate analysis and paired sample t tests to identify significant changes between pre and post.

Results
Preliminary results show very high levels of satisfaction with the 6 month program and impacts on participants similar to the 9 month program. There is evidence of personal growth on many domains in the 6 month participants including language learning, cultural diversity, citizenship and ecological citizenship objectives.

Conclusions
Evaluation is an effective tool to improve already successful programs and to optimize program interventions to produce the most cost effective results.

Keywords: Youth volunteer program; Cost effectiveness; Impact analysis; Evaluation methods

Evidence on regional inequalities in social capital formation among children in Italy. Can policy design improve children's life?

M. Spampinato
Evaluation Unit of Public Investment

Do children in Southern Italy spend too much time alone? Do they have less contact with their peers than children in the Centre and the North? Do they have less choice about where they spend their time? Do the absence of social capital lower their school performance?

Using data from two surveys by Istat (Italian National Statistical Institute) I found evidence that an asymmetric distribution of social capital between North and South affects not just the adults but, as should be expected, children. I consider both horizontal (referring to relations with peers) and vertical (referring to relations with the family) social capital. As a proxy of horizontal social capital, I used variables which show opportunities for children to meet their peers. I linked these variable with others that indicate presence/absence of private and public spaces for children: dimension of the family house in the Center North, presence of public parks (both in South and North), children’s behavior regarding reading, computer use, participation in political or voluntary activities, and children’s and teen agers playing “too much” with their parents instead of playing with peers.

I found a percentage of twenty to thirty percent of children with very poor social capital in Regions like Sicily, Campania, Apulia and Calabria. Empirical findings support the thesis that strong gaps in social capital for adults derive from problems in formation of social capital starting from childhood (path dependency).

In the graph below I report the “logit” coefficient of two multinomial models, one for the Center-North, the other for Southern regions. I used the models to estimate the probability that children grow with a strong deficit in relations with their peers, developing, thus, lower social capital.

Keywords: Social capital; Children; Regional inequalities; Human capital; Education
The experience of Sure Start: The challenges to effective outcome evaluation of a national, multi-agency initiative

N. Lloyd, L. Harrington
NLH Partnership Ltd, Evaluation, London, United Kingdom

Background
The Sure Start initiative, introduced by the UK Government in the late 1990s, is a national flagship initiative central to the Government’s efforts to improve the life chances of children living in the country’s most socially disadvantaged areas. It is an example of an ambitious, large-scale, multi-agency, area-based initiative, but despite large amounts of public expenditure, there has been an ongoing struggle to assess Sure Start’s impacts.

Purpose
This paper uses the example of Sure Start to explore key reasons why large-scale, complex, national initiatives often fail to adequately evidence the impact of their work. It provides guidance for ensuring that the impacts of such initiatives are appropriately captured.

Context
A number of recent developments have meant that effective evaluation of the impacts of publicly-funded Government initiatives has become a key priority. In particular, the recent change in UK Government, and announcement of extensive public sector spending cuts, makes the need for impact evidence more acute. In this context, what should Sure Start and similar initiatives do in order to better assess their effectiveness?

The key issues
The paper explores the key factors that have made it difficult for the impacts of Sure Start to be assessed effectively. These include a range of structural, cultural, methodological and practical factors, including factors inherent in the initiatives themselves, the cultural beliefs and practices of different stakeholders, and practical issues related to the collection and integration of evaluation evidence.

Implications for policy and practice
The paper draws together key learning on the difficulties of evaluating the impacts of Sure Start and similar initiatives. It offers practical solutions based on the authors’ extensive experience of the Sure Start initiative, and argues that more effective impact and outcome evaluation is possible, if the factors outlined in this paper are adequately addressed.

Keywords: Social policy; Outcome evaluation; UK
Evaluation of natural disasters

**The Pan-American Health Organization Health Sector Self-Assessment Tool for Disaster Risk Reduction – An Evaluation Tool for Policymaking and Programming**

E. Green, M. Zacarrelli

1 Le Groupe-conseil baastel ltée, M&E and DRM, Gatineau (Quebec), Canada
2 Pan-American Health Organization, Caribbean Program Coordination, Washington, USA

**Background and Objectives**

Health Disaster Coordinators (HDCs) in Ministries of Health (MoH) in the Caribbean, decided in 2009 to develop a Health Sector (HS) Self-Assessment Tool for Disaster Risk Reduction (DRR) with support from the Pan-American Health Organization/World Health Organization (PAHO/WHO) to undertake a comprehensive assessment of disaster mitigation and preparedness in the sector. The Tool is for sectoral policy and programming, aiding HDCs and MoHs in identifying strengths, gaps and lessons and determining priorities for health sector programs and policies, and is a monitoring and evaluation tool.

**Methods**

The construct of the Tool followed a methodology of first determining the components of mitigation (e.g. risk identification) and preparedness (e.g. governance, planning, resources) from an HS perspective, leading to a selection of quantitative and qualitative indicators, which together provide a comprehensive description of mitigation and preparedness in the HS. The data needed for the indicators then generated the set of questions, which comprise the Tool.

The Tool was developed as a Self-Assessment Tool, intended to be utilized internally by sectoral managers and policymakers to determine priorities for a national HS DRR programme or policy. The Tool was also developed as a monitoring and evaluation tool.

**Results and Application to Practice**

Baastel developed the Tool through a consultative process with HDCs and other national/sectoral officials in the Caribbean for PAHO/WHO and applied the Tool first in 3 pilot countries in 2009, and is currently undertaking application of the Tool in 5 more countries in 2010. HDCs and MOH agreed that the Tool and its application are of great value at the sectoral policy and programing level. The Tool fills a gap in terms of providing sectoral policymakers and managers with an instrument to assess where the sector is in terms of DRR and how it can move forward.

**Keywords:** Disaster Risk Reduction; Health Sector; Monitoring and Evaluation Tool; Evaluation in Developing and Transition Economies

---

**Evaluation of UNICEF’s Response to the 2004 Indian Ocean Tsunami: Lessons Learned from an Impact-Focused Evaluation**

K. Belbase

UNICEF, Evaluation Office, New York, USA

UNICEF’s Evaluation Office has recently completed an evaluation which assesses the outcomes and impacts of UNICEF response (between 2005 and 2008) to the tsunami in Indonesia, Sri Lanka and Maldives. During the inception phase, it was realised that a rigorous impact evaluation would not be possible due to the lack of baseline data, timing issues and resource availability. Nevertheless, the evaluation was conducted by keeping a focus on outcomes and impacts and examining UNICEF’s role and contribution in education, child protection, health and nutrition and water and sanitation, sectors where UNICEF was either a lead agency or had made significant investments. The experience has generated a number of lessons that can be useful in organising impact-focused evaluations in similar complex post-disaster contexts:

- Impact evaluations should be considered prospectively (planned and resourced adequately from an early stage) and focus both on learning and results (accountability).
- Timing of impact evaluation is critical: Tsunami experience suggests that the recovery phase of a major humanitarian response can take 4–5 years. It is important to decide well in advance when an impact evaluation will be carried out and to determine its scope. Evaluating too many sectors simultaneously can be a complex undertaking.
- Counterfactual/Comparison Group: Many possibilities exist and methods vary. Design should be dictated by the evaluation questions that need to be answered and practicalities involved in a certain context. Analysis of factual should not be undermined.
- Attribution/plausible contribution: Agency-specific attribution appears can be difficult and may not be that important. Joint impact evaluation may be a more useful option.
- Team composition: Important to include sector specialists who understand impact as well as humanitarian/recovery programming specialists. Involvement of national team has many advantages (knowledge of local context, language, capacity development).
Global networking to promote the evaluation profession

In recent years there has been a remarkable growth in the number of national, regional and international associations or societies of evaluation professionals. There are now at least 102 such organizations according to the database managed by the International Organisation for Cooperation in Evaluation (IOCE). The IOCE serves as an over-all promoter of networking among such organizations and their development and capacity building. The membership of IOCE is made up of national and regional evaluation associations and societies (not individual evaluators). The EES is an important partner in the mission of IOCE.

There are other international networks of individual evaluators, including IDEAS (International Development Evaluation Association. And many websites and listservs serving our growing global community.

This panel will build on recent developments in the international evaluation community and what established organizations are doing in and with that community. The audience will then be encouraged to participate in a brainstorming session to identify, with regard to partnerships among evaluation organizations generally and through IOCE in particular:

1. What has worked?
2. What could be done better?
3. What value do organizations such as EES bring to developing countries?
4. What value is there for EES in this work? and
5. What role and responsibility should a network like IOCE have?

Keywords: Networking; Global
Monitoring and evaluation bottom up and top down

Monitoring and evaluation under the new aid approach: role and influence of Civil Society Organisations in the South

M. Gildemyn
Institute of Development Policy and Management

The wave of reforms within the Aid architecture has had important consequences for the monitoring and evaluation (M&E) systems of donors, but especially of governments in developing countries. Following the Paris Declaration in 2005 and the Accra Agenda for Action in 2008 partner countries, supported by donors, have embarked on a series of reforms to make national M&E systems results-oriented, transparent and participatory. Donors, on the other hand, are trying to align their own M&E system with the partner country's systems. These reforms should enable country-led M&E systems to better perform their dual function: strengthening (downwards) accountability to ensure the implementation of (pro-poor) policies and, second, providing feedback into the policy cycle to improve programmes and policies over time. Although Civil Society Organisations (CSOs) often figure within the national M&E plans, little systematic research is available about the role and influence of CSOs within these M&E systems or through independent M&E.

This paper will explore how CSOs are able to create and gather knowledge to strengthen accountability and/or provide feedback into the policy cycle either by participating within the national M&E system or through independent M&E. The potential role and influence of CSOs will first be discussed by looking at several theories and concepts from a variety of disciplines. Theories around accountability give insights about the strategies, opportunities and limits of CSOs in strengthening downwards accountability; the literature around evaluation use and influence, and the knowledge utilization research in general, provides tools to understand the different ways in which evaluation can (or not) influence policy-makers; theories from public policy analysis give inside into the dynamics behind policy making and change. Preliminary findings from research undertaken on M&E activities of CSOs active within the health sector in Ghana will also be presented.

Keywords: Civil Society Organisations; M&E; Paris Declaration; Ghana

Access, Efficiency, and Equity Considerations of the WorldBank’s African Virtual Development Project: A Case Study of Kenya

G. Nyabade
Go Fishnet Youth Project, Community Development, Kisumu, Kenya

This paper seeks to accomplish four objectives. The first is to examine the historical development of AVU: an information technology (IT) distance learning project initiated by the World Bank in Kenya and now in over thirty sub-saharan African Countries. The second is to evaluate access, efficiency and equity implications of the World Bank's education policies in Africa with special attention to Kenya. The third is to critically examine the success and failure of the AVU as a crossborder delivery method of higher education. The forth is to provide a model for design and successful delivery of international virtual learning programmes. To achieve these objectives, a critical analysis of relevant literature was conducted. The literature search include computerized search of accessible and available material on AVU and distance education in Africa, manual search of existing literature, and communication with key subject matter experts to locate published and unpublished studies. The results of study show that AVUs design and implementation in Africa was flawed like many other World Bank education projects. While distance learning should lower costs of higher education, this is not the case in Kenya. A model for design and successful delivery of online partnership programmes is recommended.
**Theories and visions II**

**Quick, useful and dirty? – Embedding reflective practice in programme implementation**

*M. Schroll*

Care International Zambia, Monitoring and Evaluation, Lusaka, Zambia

CARE International in Zambia is currently rolling out a programme quality initiative which is managed by the organization’s Monitoring, Evaluation and Learning Unit. The mandate of the unit is to provide independent programmatic oversight and capacity building in monitoring and evaluation.

Evidence of around 20 different projects show an unbalanced relationship between mid-term and final evaluations with more efforts and funds reserved for the latter (ca. 1:2). However, this paper argues for an increased attention to mid-term evaluations in project/programme settings in the form of In-Action Reviews (IAR).

IARs are utility-driven, simple, actively involve project management and partners, include analysis of project data and are facilitated by an independent entity. The purpose is to initiate and facilitate a reflective process leading to ownership of results and immediate improvements thereby embedding evaluative practice in programme implementation. The reviews focus on efficiency and effectiveness and incorporate an exploratory view toward impact and sustainability. The recommendation section is replaced by options and no formal standard report is issued. Instead, findings are summarized in a presentation – using visual aids such as graphs and photos – and emphasis is put on discussing results with staff and partners. The existence of the Monitoring, Evaluation and Learning Unit as an internal service provider ensures transparency, feasibility, timeliness and independence of the process as well as follow-up.

The paper briefly summarizes an IAR process and its effects and discusses the pros and cons in the light of aspects of evaluation quality such as accuracy, utility and independence.

**Keywords:** Utility; Feasibility; Reflective practice; Mid-term evaluation

---

**Survival of Public organizations in the new era of planning**

*A. Caetano*, T. Costa, N. Cavaco

1 Ministério da Administração interna, Secretaria geral, Lisboa, Portugal

2 Leadership, Leadership, Lisboa, Portugal

Managing the Portuguese public services is, increasingly, a complex activity where ranges of different types of organization are involved in the delivery of public services. Public services managers have had to develop new skills and adopt new perspectives as the boundaries between public, private and voluntary sector organizations become blurred. The management task becomes one managing ambiguity in an ever-changing world. At the same time, however, there is a certain timeless to any debate concerning the management of public policies and managers will need to acknowledge the continuing relevance of traditions and the enduring nature of themes of accountability, responsibility, acting in public interest, integrity, probity and responding to citizens, clients and customers. At the same time they create double standards thus for every management system in organizations, there is a shadow system that is used to run the official system.

All of this increases complexity of the system and increases the necessity of using new kinds of process management were new knowledge and new Skills are fundamental.

We are going to try to tackle the following questions: How can intuition be use in process of planning? How can we implement process management? What kind of knowledge is necessary and how can we implement processes of qualifications that produce organisational development? Can evaluations designs and methods originating in high trust environments simply be used in low trust environments? Will institutions, agencies and societies that manage resource driven and resource oriented interventions for their own benefit simply submit to externally imposed “objective” evaluations or will they try to “manage” evaluations as well?

**Keywords:** Evaluation public services; Planning
Is there a place for holistic evaluation in urban planning?

L. Horelli, S. Wallin
Centre for Urban and Regional Studies, Aalto University, Helsinki, Finland

During its long history of different waves evaluation has become part of normal praxis in most areas of public life. This is not, however, the case with urban and spatial planning which define the long-term conditions and placing of activities that belong to different sectors of policy and administration. In many European countries, Finland included, evaluation in urban planning is only restricted to the phase of actual planning process (The EU compendium of spatial planning systems 1997). In addition, recent compilations of critical essays in planning theory totally ignore evaluation as a major tool for developing planning theory and praxis, as well as better environments (Hillier and Healey, 2008). Even the publications that deal with the subject, such as the “Evaluation in planning” restricts the target of evaluation to ex-ante impact assessments (Alexander, 2006). The same is true for the few examples of evidence-based practice in planning (Krizek et al. 2009). There are, however, a few solitary articles which claim that it is important also to assess the institutional conditions and ways of implementing the plans (Oliveira & Pinho 2009; Talen 1997).

We argue that the expansion of evaluation from the ex-ante impact assessments of the planning phase to a holistic evaluation also of the implementation and post-occupancy phases, which provide feedback to the knowledge base of strategic planning, might improve both planning theory, praxis and environmental settings. The aim of the paper is to present an analysis of the literature on evaluation in planning and to make a proposal for a model of holistic evaluation in planning based on the Finnish system of planning as an example.

Keywords: Urban planning; Ex-ante evaluation; Post occupancy evaluation; Impact assessment; Finnish planning system
Methodological processes

New methods? Perspectives on changing tasks in the field of evaluation

M. Wikner¹, R. Ingrid²

¹ Ramböll Management Consulting, Welfare Analysis, Stockholm, Sweden
² Ramböll Management Consulting, Growth Analysis, Stockholm, Sweden

Background

In both literature and in the practical evaluation work, we experience a trend towards evaluation commissioners involving evaluators as integral to the whole policy process (see among others “From studies to streams: managing evaluative systems”, Rist & Stame, 2006). Traditionally, the role of the evaluator has been tied to the later steps of the policy process – e.g. retrospective evaluations. Today, the evaluator tends to be active in virtually all phases of the policy process; problem identification/needs assessment, recommendations for solutions, tailoring of intervention as well as implementation and evaluation of the same.

While the evaluator “moves upstream” in the policy process, the evaluation assignment changes. It is our experience evaluators often lack a strategy to handle the broadened assignment. There is an uncertainty of how the new role should be understood, and a lack of suitable methods and tools. Evaluators depend on sound and systematic methods to deliver high quality evaluations. But have evaluation processes and methods developed in accordance with these broadened tasks? Is there a risk our methods erode when applied to new parts of the policy process? A move upstream in the policy process demands a widened range of methods and an ability to adapt to a moving target during the problem identification. At the same time the implementation phase demands hands on tools for organizational development.

Objectives

Our paper aims to give an overview of implications of a broadened assignment on choice of methods and tools to provide high quality evaluation services in all steps in the policy process.

Methods

Analysis of process and content of four reference assignments from the Swedish context

Results

When the scope of an evaluation is the entire policy process, there is a need for a broadened discussion on what methods are relevant during different phases of the policy process.

Keywords: Policy; Methods; Changing evaluation tasks

Measuring international collaboration, challenges in methodology of evaluation.

A. Kalyta

Science and Technology Center in Ukraine, Performance/Public Outreach Department, Kiev, Ukraine

Background

The Science and Technology Center in Ukraine (STCU) is an intergovernmental organization that assists former WMD experts from ex-USSR in the transition to self-supporting, peaceful activities in the international science communities, using the best professional practices. One of the tools for reaching its objectives is sponsoring the events, aimed at the establishing international collaboration with the scientists from western countries.

Objective of the presentation is to overview the methodology, applied for “measuring the collaboration” and its certain challenges.

Methodology

For defining the effectiveness of participation of beneficiaries of STCU grants in conferences, seminars, exhibitions evaluation research was conducted since 2006. The set of indexes was applied: indexes, characterizing joint activity, established by western and ex-USSR scientists (quantity of joint projects, publications, internships etc); indexes, showing commercialization of scientific developments (contracts, licenses sold etc), usage of the information, obtained during the event. Also qualitative information about the results (stage of collaboration or negotiations) was gathered.

Results

Certain challenges have appeared during the evaluation study. Some of them together with solutions taken can be presented to the audience and discussed, e.g.:

• difficulties in dealing with qualitative information, while comparing the results of the events (what is better – one contract or five joint articles?);
• long term needed for obtaining the tangible result and necessity to report to donors
• necessity to show the modest effectiveness of some programs and the risk to have the financing cut;
• difficulty to evaluate the volume and usefulness of scientific and marketing information, obtained by the scientists during the events;
• transparency of the evaluation and non-disclosure policy;

Application to practice

The issues are often debatable and can be of interest for representatives of donor organizations and evaluation practitioners. The overview of the literature on such dilemmas and solutions taken in STCU will be elucidated.

Keywords: Methodology; Impact assessment; Qualitative data analysis; Donor organizations
Learning to self-evaluate at University

S. Pastore
University of Bari

In the light of the current changes which have affected the University system so far, a new evaluation practice is widely standing out. It mainly aims to embrace the students’ opinions about the formative services provided/received within a University degree course.

Differently, evaluation for University examinations does not attract the scholars’ current research interests. Nowadays the learning measurement and testing frequently seem to be practised more as a mere bureaucratic tool than as a pedagogical evaluative method because the shadow of sanction hangs over those educational practices which should lead the teacher to reflect on the student’s un/success and on his/her learning enhancement.

Despite increasing interest in formative evaluation, it remains largely an enigma: an examination of the literature and an attempt to establish a clear definition can lead to a range of complex problems.

The vagueness of operational definition has contributed to the paucity of research in the area and to lack of empirical study that could help identify so-called best practice in evaluation.

Without a proper understanding of the meaning of formative evaluation, it is clear that any research conducted will make a poor contribute. Moreover, the overlap of terms and concepts and the ambiguity of terminology are a good indicators of how necessary it is to identify what formative evaluation is and what action may be implemented.

I will present findings of a case study conducted at the Faculty of Education, University of Bari. The paper will critically examine common educational belief and its related problems – particularly linguistic – in order to show that formative evaluation, even on university setting, can prove an effective learning experience.

From the analysis it is clear that we can consider evaluation in much broader sense than the traditional practices of grading – e.g., the traditional paper-pencil tests or oral exams.

Keywords: Evaluation; Learning; Reflection
Disentangling complexities

Reaching the parts other evaluations cannot reach? Evaluating complex and unfolding interventions in real time

T. Ling
RAND Europe, Director Evaluation and Audit, Cambridge, United Kingdom

Evaluating efforts which are designed to adapt and learn, which respond to a changing environment, and where the aims might shift over time can be seen as an evaluator’s nightmare. Practical and methodologically robust guidance on how to do this is limited. Furthermore it is often expected that under such circumstances evaluations should be delivered in real time and provide both formative and summative roles. Evaluators then face questions of how to provide independent judgements in real time that are both helpful and avoid being captured. Evaluation should not stay in the areas where it is safest and easiest. An approach aiming to achieve certainty about attribution is likely to provide neither effective learning nor identify accurately the benefits achieved. Using examples from recent RAND Europe evaluations in the areas of health, Impact Assessments, and regulation, Ling will illustrate how an alternative approach can work. The approach draws upon the contribution approach of John Mayne, advocates focusing scarce evaluation resources where uncertainties are greatest and most significant, designing learning outcomes and processes from the outset, and ensuring that accountability to stakeholders is strengthened. In brief, this focuses on Investigating contribution stories; Reducing key uncertainties; Improving learning; and Supporting accountability (IRIS). This paper is intended as a basis for discussing this vexed set of questions.

Keywords: Evaluation; Methods; Contribution; Learning; Accountability

Independence and Evaluating Complex Policies

P. Casavola 1, L. Tagle 2
1 Independent Expert, Independent Expert, Rome, Italy
2 Evaluation Unit, Department for Development and Social Cohesion, Rome, Italy

Independence is one of the pillars of evaluation credibility. It is a matter of both ethics and institutional arrangements for requesting, financing, conducting, and disseminating an evaluation. In the world of practice, they are strictly interrelated – for no institutional arrangement can make up for unethical behavior and, yet, unethical behavior may be at least partially deterred, detected, or made useless by accurate institutional constructions.

When evaluation deals with relatively simpler or impersonal policy mechanisms, much can be gained by separating evaluation research projects from policy making, relying on the need of building a credible reputation by independent experts. However, increasingly evaluation questions—even when apparently focused—relate to complex policies and programs implemented in intricate contexts where separating evaluation research from policy making proves both more difficult and less effective. Moreover, when policies are inherently complex, a judgment on the policy is perceived as a judgment on the policy maker and, since there is more at stake for the policy maker, independence is more at risk.

The ultimate paragon of independence in the world of evaluation is traditionally identified as an evaluation unit responsible to a collective, possibly elected, body, which, in turn, is autonomous and sovereign. In the European context, though, such paragons are precisely few. Most often, it is financing, coordinating, programming, and (much more rarely) implementing entities which request evaluations. What does it mean, in this context, to build an institutional arrangement conducive to independence? How (and whether) should institutional arrangements change across the divide between internal and external evaluations? How does this change when considering that most interventions are the responsibility of complex groupings of organizations?

This paper focuses on institutional settings and explores merits, trade-offs, and opportunities of alternative ways of ensuring independence. It will use the experience gathered within regional policy in Italy.

Keywords: Independence; Complex policy; Institutional arrangements for evaluation
Knocking on heaven’s door – outcomes measurement arrives at church

V. Denvall, S. Linde
Lund University, School of Social Work, Lund, Sweden

Verner Denvall & Stig Linde:

Knocking on heaven’s door – outcomes measurement arrives at church

Ideas and concepts of how to evaluate travel around the globe as the diffusion of evaluation proceeds. Studies of how evaluation is disseminated, diffused and implemented are important if we wish to keep ourselves updated with the change-patterns of the evaluation-landscape. We have a limited knowledge about the procedures of uptake, change and use of evaluation in different contexts. The proposed paper session challenges the idea “through increased transparency and pluralistic approaches, evaluation also triggers pressures for improved performance and it restrains abuses of power” (Cited from the conference proposal).

In this session we report from the uptake of accountability in three congregations in The Church of Sweden. The church is of specific interest when focusing evaluation since it for almost 2000 years has used internal ways of assessing accountability but now must respond to the globalizing influence of evaluation. Based upon an analytical framework from organizational theory we will show how the same concept of outcomes measurement have been adopted, rejected or transformed depending on the local policies and how central policies have failed in implementation.

In the study presented we have looked at an organization with a history of traditions and legitimacy and a successful audit of its own and how it responds to modern concepts of administration where evaluation plays an important role. Non-profit organizations have to balance between spreading values and to produce with efficiency and control. We will present how modern policies of administration and ways to implement principles of evaluation and quality assurance meet longstanding traditions of autonomy. This session contributes to a better understanding of the field of evaluation by discussing the transformation that such modern concepts of organizing undergo in the journey between and within organizations.

Keywords: Transformation; Church; Diffusion; Local policies; Outcomes measurement
Civil societies organizations promoting evaluation

A civil society-initiated, legally binding evaluation system for local government: evidence from São Paulo

F. Rigout
Plan Políticas Públicas, Research and Consultancy, São Paulo, Brazil

This paper tracks the history of an evaluation framework originating entirely in organized civil society. In 2007, the non-partisan umbrella organization Nossa São Paulo (NSP), in the city of São Paulo, Brazil, began a lobbying campaign to pass a measure that would mandate every entering administration to agree to a set of results targets to which it would be held legally accountable at the end of the four-year term. Named Agenda 2012, the monitoring and evaluation system is now in full effect, encompassing every department of the current administration, totaling 223 results indicators along six axes: Human Rights, Sustainability, Creativity, Opportunities, Efficiency and Social Integration.

My research is based on two perspectives. As an outside observer, I discuss the instances of conflict and mutual learning between City Hall and NSP when specific targets had to be agreed upon. This relationship generated strains regarding the use of evaluation data, making evident the tensions between the dire need for long-term planning in a city of 10 million people, and the administration’s initial project of embracing accountability largely as a means to publicize its four-year agenda. As an insider, I recount my experience as the head M&E consultant at the city’s Housing Department, where we tried to implement results-based management amidst the attributions-based routine of technocrats, whose unfavorable view of evaluation began to change as the Agenda 2012 took hold.

Funded by local businesspeople, NSP brings together more than 600 civil society organizations and community associations. It has “no president and no board of directors, began as a social network and expanded as such”. I claim that their organizational model and political strategy are effective means to boosting public sector accountability in democratic, middle-to-high income developing societies, where government legitimacy is increasingly dependent on the attainment of measurable results.

Keywords: Civil Society; Developing Democracies; Local Government; Government Accountability; Brazil

The role of civil society in building evaluation capacity

R. Mihailestone
Pluriconsult Ltd., Romania

The existence of a professional association was considered as one of the “9 + 3 criteria” of an established evaluation culture (i.e. the institutional and political framework in which the evaluation practice is carried). To what extent the existence of such an association should necessarily indicate a fully established evaluation culture? Of course this criterion should be judged in the broad institutional and political context evaluation operates as this has an important influence on the mission, strategy and action plan of the respective professional association. But what could be the contribution of such an association to the development of the national evaluation system? Is this only limited to establishing norms and code of ethics for the practice or could it be more, such as advocating for strengthening the evaluation demand or increasing the role and the use of evaluation for improving governance? And how could this be done in practice? With what effects? The paper will address the questions above by making an analysis of the existing practices among the European evaluation associations/societies.

Keywords: Evaluation capacity building; The role of evaluation societies
Evaluation in HIV AIDS programmes

Effectiveness Evaluation of HIV/AIDS Prevention Programs for Men who have Sex with Men of Non Governmental Organizations in Brazil

M. Cruz, E. Moreira dos Santos, S. B. Santos, P. V. Decotelli, R. Torres, A. Reis
Fundacao Oswaldo Cruz/ENSP, DENSFLAES, Rio de Janeiro, Brazil

Background
The HIV/AIDS prevention programs of Non-Governmental Organizations (NGO) in Brazil have focused on men who have sex with men (MSM), particularly in young and most vulnerable group. However, there are few initiatives to evaluate the effectiveness of such programs and verify if they have contributed to sexual behavior and attitude changes in promoting safe sex life and healthy.

Objectives
A case control study to evaluate the effectiveness of HIV/AIDS prevention programs for MSM by NGO in Brazil and to determine favorable aspects and barriers that affect these actions.

Methods
Four NGO which had successful preventive actions were selected for further study in order to explain the relationship between intervention and actions of prevention among cases and controls and finally determining the effectiveness based on evaluation criteria established. Data was collected through documental analysis and interviews with managers and users.

Results
Show that actions developed by NGO do not necessarily follow what is in the scope of project. Preventive actions are more frequent in areas of interaction between MSM and emphasis on distribution of condoms and educational materials. It has been seen greater involvement prevention among adult MSM from more favorable socioeconomic conditions, while younger MSM are at higher individual and social vulnerability.

Conclusion
Most NGO don’t have records nor monitor their actions, highlighting the need to implement a system of monitoring and evaluation. It is important to develop different strategies to face the methodological challenge of addressing research control subjects.

Keywords: Effectiveness evaluation; MSM; Prevention programs; HIV/AIDS; NGO

Implementation Evaluation of HIV/AIDS Legal Assistance Services of NGOs in Brazil

P. V. Decotelli, M. Cruz, M. Ventura
1 Fundacao Oswaldo Cruz/ENSP, DENSFLAES, Rio de Janeiro, Brazil
2 Fundacao Oswaldo Cruz/ENSFE Ensp, Rio de Janeiro, Brazil

Background
Brazilian efforts in the HIV/AIDS epidemic are, since its beginning, based on human rights (HR) with strong participation of Non Governmental Organizations (NGOs) and society. Legal services have increased through national policy and have been responsible for orienting, educating, solving conflicts and building capacity related to HR and HIV/AIDS for people living with HIV/AIDS (PLWHA) in Brazil, especially among more vulnerable population and community leaders.

Objectives
The project aims to understand and explain the implementation of HIV/AIDS legal assistance services of NGOs in Brazil. Analyzing the M&E Human Rights Violation Database of the National Program (NP), official documents and interviewing representatives of the legal services and of the NP allow describing the picture of promoting and protecting human rights for PLWHA in Brazil.

Results
Legal assistance services provided by the HIV/AIDS NGOs in Brazil and financially and technically supported by the NP are valuable tools of promoting human rights in the country and building human rights capacity. However, reduced budget and personnel affect the continuity and enlargement of their activities. Besides that, they are not able to reach PLWHA outside major cities or that are not participant of a good network of relations. Therefore, a reduced number of PLWHA who are in more vulnerable situation reach the legal assistance services and have their rights realized.

Application to practice
There must be further studies to evaluate public health and human rights programs, establishing a better framework. Articulation of the NGO services with the ones already provided by the government to reach a larger number of people and continue the process of HR promotion and protection for PLWH in Brazil. It is also necessary to improve strategies of promoting the legal services in the media, in communities and in health services.

Keywords: Evaluation; Human rights; HIV/AIDS; Legal service; Brazil
Evaluation and gender

Evaluation of a Gender Health Policy framed in the Women’s Health Plan of the Autonomous Community of Cantabria in Spain

A. Cabria, M. Pla, R. Crespo

1 Fundación Marqués de Valdecilla, Observatorio de Salud Pública de Cantabria, Liencres, Spain
2 Universidad Autónoma de Barcelona – Fundación Robert, Cátedra de Investigación Cualitativa, Barcelona, Spain

Background
Cantabria integrated in its first Women’s Health Plan 2004–2007 a new gender health policy that went beyond the scope of the plan to challenge the entire health system. This multi-method comprehensive participatory evaluation of the policy and the plan is the first of its kind in the region and the country.

Objectives
1) Evaluate the conceptual framework of the policy, its validity, and significance; the process regarding mainstreaming, efficiency, scope, and access; and the results with respect to its effectiveness.
2) Transfer knowledge and practices among stakeholders.
3) Recommend changes for improvement.
4) Contribute to generate an evaluation culture.

Methods
1) Analysis of available secondary information sources and databases;
2) Stakeholders analysis;
3) In-depth interviews and discussion groups;
4) Gender and health survey to health professionals and policy makers;
5) Consensus workshop with stakeholders.

Subjects are the stakeholders involved, including policy makers, health managers and providers, patients associations and users. Sampling was carried out by theoretical representation with maximum variation.

Results
Whereas the introduction of this gender health policy represented an opportunity for the health system to address gender health inequalities, its expected goals and ambiguity exceeded the scope and potential of the plan. Nonetheless, 13 programs and protocols of the Plan integrated a gender perspective. Despite the providers’ low level of understanding regarding the meaning and implications of using a gender lens in their clinical practice, and the resistance to apply it, some policy makers, health providers and researchers are beginning to take a gender approach.

Conclusions
Regardless of its limitations, this initiative represents today one of the few gender health policies in the country to address gender health inequalities. This participatory evaluation is contributing to give visibility and relevance to the policy as well as credibility to the practice of evaluating for equity and accountability.

Keywords: Gender health policy; Participatory evaluation; Generating an evaluation culture

Towards a gender-sensitive evaluation? Practices and challenges in international development evaluation

J. Espinosa
Universidad Complutense de Madrid, Facultad de Ciencias Políticas y Sociología, Madrid, Spain

Gender equality has been introduced into the international development evaluation not long ago. The progress of the evaluation discipline itself, the advance towards a more multidimensional concept of development and, as a result, the incorporation of gender issues into international development agenda have caused that the vast majority of evaluation practices regarding to the gender equality were found in the last two decades.

During these twenty years, the proportion of evaluations that focuses on analyzing gender issues has increased gradually, despite the growing conservative political atmosphere we are witnessing and the difficulties this political agenda is placing on the mainstreaming of gender. For this reason, evaluation is increasingly becoming a niche for dealing with gender issues, and improving gender equality policies and accountability, even though the link between gender and evaluation has yet to be analyzed in depth.

In this paper, we examine this link by both discussing several evaluation experiences in analyzing gender issues and describing the main challenges of carrying out gender-sensitive evaluations.

Keywords: Gender-sensitive evaluation; International development evaluation
Missing links: engendering evaluation criteria

P. Alvarez
INFOPOLIS, Evaluation

Most studies and meta-evaluations find a low-key approach to gender in evaluation and conclude that the gender dimension is not yet integrated into the evaluation criteria and procedures. Reasons for that range from lack of instructions and tools for programme management staff and expert evaluators, as regards the integration of gender in the evaluation process, to lack of gender knowledge or the perception that gender issues are irrelevant or even contradictory to an objective assessment of an evaluation quality.

Engendering evaluation practice requires, among other actions, to define common approaches, share good practice and propose innovative ideas. Practical tools need to be produced to help integrate gender equality and women’s empowerment into all aspects of evaluation. An important part of this integration entails mainstreaming gender into standard evaluation criteria.

While much effort has been conducted to ensure that equal opportunities and gender be integrated into data collection methods, it is not sufficient in itself to increase the participation of women or the amount of sex-disaggregated data. This paper explores the opportunities and challenges to include gender equality within evaluation criteria by both engendering standard evaluation criteria as well as adding alternative criteria to the existing ones.

Introduction of this complexity would have significant implications in terms of increased quality in evaluation practice and utility, development of specific gender competence of expert evaluators and enhanced practical knowledge about gender dynamics and successful strategies conducive to a more egalitarian society.

Gendering Evaluation: What does it mean and imply?

M. Bustelo
Universidad Complutense de Madrid. Facultad de CC. Políticas y Sociología, Dpto. C. Política y de la Admón. II, Pozuelo de Alarcon (Madrid), Spain

In this positional paper I aim at exploring how evaluation and gender interplay in regards to public policies. Evaluation might be seen as part of the policy making process, and as a policy tool by itself. As a policy tool, the evaluation might be especially fruitful for capturing the important changes and shifts the policies which aim at fighting against gender inequality (i.e. gender policies) have experienced during the last two decades, for improving those policies, as well as for answering to what extent these policies are successful. As part of the policy making process, and following the aim of the internationally recognized gender mainstreaming strategy, evaluation is an important part to be conducted under a gender – and human rights and other inequalities – perspective; however, the implications of what it means to do evaluation from a gender perspective still remain to be thoughtfully explored.

Through this paper, I first intend to clarify the various ways in which evaluation and gender interplays at answering questions covering that double perspective already mentioned. On one hand, I distinguish between two important aspects to take into account in the evaluation of gender policies nowadays: the evaluation of specific gender policies or gender components within broader policies; and the evaluation of the gender mainstreaming strategy by itself. On the other hand, I want to explore the role gender potentially plays as part of the evaluation process, in any kinds of public policies —and not only gender ones—. What does it mean doing an evaluation from a gender perspective as the gender mainstreaming strategy suggests? What does it imply? Is it very different from a ‘regular’ mainstream evaluation? If we mainstream a gender perspective, it is also possible to mainstream other inequalities and a human rights perspective?

Keywords: Gender; Evaluation from a gender perspective; Gender mainstreaming evaluation; Gender policies evaluation
Transparency

Are we equipped to deal with the human and political factors in our evaluations? A few lessons based on international experience

M. Gervais
Université Laval/ICSS de la Vieux-Capitale, Médecine sociale et préventive, Québec, Canada

For evaluators, from the greenest beginners in their first training program to seasoned professionals well into their practice, the emphasis is often placed on developing technical skills (design, data collection, analysis, interpretation, and reporting) and management skills (evaluation process management, budgeting, resource coordination, and supervision). Little time and few learning opportunities are devoted to “softer” skills, such as reflective skills (norms and standards of practice, meta-evaluation), interpersonal skills (communication, negotiation, conflict resolution, collaboration, transcultural skills), and situational skills (analysis and accommodation of specific interests, contextual circumstances and issues involved in the evaluation process).

This presentation will draw on the experience of evaluations conducted on several continents (mainly North America and Africa), presenting varying levels of complexity and difficulty. It will show how insufficient attention to these “softer” skills leads evaluators to make non-optimal decisions and can interfere with, even shackle, the evaluation process, to the point of jeopardizing expected success in addition to considerably limiting the expected influence of the evaluation on managers’ subsequent decision-making processes.

What can we take away from this? How could certain risks, attitudes, or events relating to human and political factors be identified sooner, dealt with, and monitored in such a way as to minimize their effects on the rest of the evaluation? What tools are available to guide evaluators in such situations? What strategies might they establish to better manage the conduct and success of an evaluation? What positions might evaluators take?

This presentation will allow us to learn from these experiences and explore a few concrete ways to better support evaluators dealing with these issues and better equip them for developing best practices.

Keywords: Human and political factors; Reflective and situational practice; Risk analysis; Decision-making process; International experiences

Policy commitment and evaluation in Slovakia: Evidence-based analysis of procurement decisions made in public sector evaluations

D. Gombitova, D. Aligner, A. Frohlich
Slovak evaluation society, Slovak evaluation society, Bratislava, Slovak Republic

Slovakia is still facing challenges in relation to evaluation culture in general and the proper management and use of evaluation findings in particular. The uneven commitment of national policy and public administration to evaluations can be observed in the way evaluation assignments are prepared, organised and procured.

The evaluation itself should serve as a tool for enhancing transparency and supporting the accountability of public financial management. To fulfil this purpose, it is therefore essential that the process is managed properly from the beginning, and that the evaluation service is procured in an unbiased and transparent way.

This Paper presents the results of a recent survey whose aim was to assess the public procurement of evaluations in Slovakia. Data compiled at the level of the central government bodies for the period 2006 – 2009 has been systematically analysed. The survey indicates that the risk of biased procurement decisions in the area of public evaluations is substantial. The interpretation of available data suggests that there is a real risk of an inefficient and non-effective use of public funds spent on evaluations.

Consequently, the evaluation function, serving as an impartial and evidence-based feedback and supporting tool for public policy decision makers, is not perceived as being helpful or important. Evaluation is not generally accepted as an integrated part of the policy cycle, but is seen as a matter of formal compliance. The use of evaluation is limited and its role distorted. There appears to be a wide-spread lack of administrative knowledge and experience about the benefits of evaluation. Moreover, national procurement procedures and guidelines for evaluation, although being compliant with the legal framework, often suffer from ethical problems which are unlikely to be resolved without demonstrated political commitment.

Keywords: Transparency; Procurement; Ethics
Self evaluation

Building capacity for mandated self-evaluation in Iceland schools through collaborative cultures and professional learning communities

P. Lisi1, S. Davidsdottir2
1Central Connecticut State University, Educational leadership, New Britain, USA
2University of Iceland, Education, Reykjavik, Iceland

Background
In 1996, legislation was passed in Iceland mandating that upper secondary schools engage in self-evaluation. Findings from a longitudinal study of school responses to the mandate underline the significant importance of leadership, goal setting, and professional development. Though not significant, a culture of collaboration is emerging as a new norm as Icelandic schools build self-evaluation systems. While research about the impact of collaborative cultures and professional learning communities (PLCs) on school improvement is compelling (Blankstein, 2004; DuFour, 2003), widespread adoption of these practices is not the norm. In an international study of teaching and learning (TALIS, 2009), Icelandic lower secondary schools ranked low in the professional collaboration index (e.g. team teaching, observing other teachers, engaging in professional learning).

Objectives
For Icelandic upper secondary schools, important questions are: 1) To what extent is a collaborative school culture becoming institutionalized? 2) What is the impact of emerging collaborative school cultures on teaching, learning, and self-evaluation? 3) What factors contribute in significant ways to the development of collaborative school cultures and PLCs?

Methods
This research is a mixed methods descriptive study in Icelandic upper secondary schools. Data were collected through a questionnaire circulated in 11 schools (250 teachers), and interviews conducted with evaluation team members in 5 schools.

Results
Findings suggest that while teachers are engaging in collaborative work (e.g. evaluation, curriculum development, action research), challenges to the process (e.g. creating opportunities for public discussion of evaluation and making time for collaboration around performance data) need to be addressed.

Conclusion/Application to practice
Building a culture of collaboration is particularly interesting in a country in which a strong tradition of independence and individual initiative has existed for centuries. Schools are recognizing collaborative cultures and PLCs as promising practices for enhancing the legitimacy and transparency of the evaluation process.

Keywords: Participatory evaluation; Building evaluation capacity; Teacher empowerment; Collaborative cultures; Professional learning communities

Strenghts and weaknesses of self-evaluation in Slovenian schools

M. Brejc, A. Juric Rajh
National School for Leadership in Education, National School for Leadership in Education, Kranj, Slovenia

In the framework of European Social Funds and Ministry of Education and Sports co-financing in years 2009–2014 national system of quality in education (NSQE, hereinafter) will be introduced for pre-primary, primary and secondary education in Slovenia. The aim is to gradually introduce the system, that will be based on school self-evaluation, external evaluation and quality indicators.

The development of the system is based on the presumption that capacity building is in the heart of school improvement and that school system can only progress by individual and system learning. Strategic changes introduced must therefore influence not only the visible characteristics of the system but also values, beliefs and expectations of people in everyday practice (Elmore 2007).

From April 2009 73 schools are being trained for self-evaluation. A training transfer model has been developed in which a head teacher and 2 teachers or other professionals from each school (school team) attend the trainings and then further introduce and implement self-evaluation processes in their own schools. During the period of 1 year schools have participated in 4 one day trainings and were introduced with topics, such as: (1) students’s achievement as a self-evaluation focus, (2) annual improvement planning, (3) defining «measures» and gathering data, (4) the role of stakeholders in self-evaluation. Between each day of training school have intermediate activities that are based on professional discussions among staff, annual planning on school and teacher level, identifying the role of primary stakeholders in self-evaluation, etc.

Presentation will focus on strengths and weaknesses of self-evaluation process as staff from 73 schools perceive them after 1 year of training and self-evaluation process, in terms of (1) importance and (2) current perception in practice. Conclusions and recommendation will be discussed for further trainings.

Keywords: School self-evaluation; Training; Perceptions
Self-evaluation in Icelandic secondary schools – influences within the classrooms

S. Davidsdottir¹, P. Lisi²
¹University of Iceland, Education, Reykjavik, Iceland
²Central Connecticut State University, Educational leadership, New Britain, USA

Background

Icelandic secondary schools were mandated in 2006 to do self-evaluations. They have met this mandate with different approaches and great differences in teacher participation. In some schools teachers own the process of evaluating schoolwork and producing developmental plans, whereas in others one or more persons are assigned the task of conducting evaluations.

Objectives

The authors were especially interested in what would be associated with teacher participation in self-evaluations, and use in classrooms.

Methods

Participants were 250 teachers from 11 Icelandic secondary schools. Their responses to questions on schoolwork were factor analyzed. A stepwise regression was used to explore the variance in self-evaluations which could be explained with the other factors.

Results

Five factors were found: leadership, professional development, school goals, collaboration on initiatives, and self-evaluation participation. Those factors that explained most variance in self-evaluation participation were: school goals, leadership and professional development, all statistically significant. Collaboration on initiatives in schoolwork did not add significantly to the explanation of variance in self-evaluation participation.

Conclusion

It seems that the most important factor influencing teacher participation in, and use of self-evaluation pertains to school goals. Specifically, teachers are most active when 1) school goals are informed by data about student achievement; and 2) the main criterion for determining progress in achievement of school goals is student achievement data.

A positive relationship with leadership was also important. Professional development should inform self-evaluation practices, which seemed to be the case here. Collaboration in general is not widespread in Icelandic schools, as was found in TALIS 2009.

Application to practice

It seems that in order to encourage active teacher participation in evaluation work, leaders should encourage strong use of student achievement data to inform development of school goals and facilitate related and powerful teacher professional development.

Keywords: Self-evaluation; School goals; Leadership; Professional development; Collaboration

Vocational skills demonstration in practice in vocational education and training in Finland

A. Räisänen, DHSc, M.Sc. Head Coordinator.
Finnish National Board of Education (Finnish Evaluation Council. The leader of this national evaluation), Helsinki, Finland

This paper presents the methodology of the national evaluation of skills demonstration in vocational education and training organized by the Finnish Education Evaluation Council. The evaluation concentrates on assessing the efficiency of the system of vocational skills demonstration in practice, and more particularly on the planning, execution and evaluation of the vocational skills demonstrations in collaboration with working life. It focuses on the evaluation of the set goals for vocational skills demonstrations, on the overall cost effects and the administrative effects of the vocational skills demonstrations. The object of evaluation is curriculum based vocational education.

The evaluation was based on development-oriented evaluation, and it was criteria based. The evaluation model used was the Common Quality Assurance Framework (CQAF) developed by the European Union for assessing the quality of vocational education. The evaluation data was collected from multiple sources and with multiple methods. The providers of education and the degree based groups described their actions related to vocational skills demonstrations, produced the quantitative data, and assessed their own action commissioned by a diverse evaluation team. The providers, educational institutions, and directors, representatives of working life, teachers and students participated in the self evaluation. The external evaluation group paid six evaluation visits. Various reports and follow-up data on the costs of education and learning outcomes provided by the National Board of Education were used as sources of information. The response rate was 96 %. The external evaluation was organized by the multiprofessional team.
Evaluation of health sector intervention

Online web based system of data management for Post Acute and Long Term Care

J. Abreu Nogueira, A. Girao, I. Guerreiro
Ministry of Health, Mission Unit for Continuous Integrated Care, Lisbon, Portugal

Background
Benchmarking and research are essential for post acute and long term care. The use of an integrated tool for bio-psychosocial evaluation implies an extensive data analysis as well as other data. Registration on paper needs further construction of data bases to analyze results.

Objectives
To develop a paper-free on-line web based system of data management for the National Network for Continuous Integrated Care (RNCCI), that allows on line registration of the evaluation made with the integrated bio-psychosocial tool, data related to patient safety and the registration of data related to referrals from hospitals and primary care and admissions to RNCCI, that allow real time results, pertinent to management and for professionals.

Methods
Software prototyping (allows continuous improvement through feed-back from users) is the basic development ongoing methodology. The web-based solution has a relational database management system to allow queries and different reports models and different access levels.

There was a focus on professional training related to the IT tool and registration. There are online reports concerning the different parameters registered.

Results
Its implementation began in the first trimester of 2008 on a voluntary basis with continuous engagement of professionals. Nowadays all the referrals are made online allowing characterization of hospital management of discharges to RNCCI, characterization of the users referred allowing monitoring of adequate referrals as well as individual parameters registered with the integrated tool for bio-psychosocial evaluation and the monitoring of the outcomes related to deliver of care.

Conclusion/Application to practice
This data management system allows having real time results at a national, regional, local, and unit level, making benchmarking possible to obtain with a large number of registries. The use of similar methodologies would contribute to research in post acute and long term care

Keywords: Online; Data management; IT; Long-term care; Evaluation

The Politics of the Centre and Periphery in Cambodia’s National Health Service and International Interventions

I. Ramage, K. H. Ramage, E. Mazard
Domino Research and Consulting, NA, Phnom Penh, Cambodia

A series of evaluations concerning the improvement of Cambodia's health sector revealed crucial factors in the structure of implementation and the politics of distribution that can determine program outcomes, regardless of policy.

In an industry dominated by statements of principle and abstract targets (such as the MDGs) this study directs us to a process-oriented view of implementation that acknowledges the political and economic interests of the actors at each level.

The results draw on a decade of experience in Cambodia, with special reference to recent evaluations measuring outcomes prior to a new round of strategic planning for Cambodia's internationally subsidized public health sector.

The data was collected in surveys of randomized populations, of beneficiaries of specific health programs, and also facilities-based investigations. This includes a survey of 926 households distributed in 44 different villages, and inspections of 20 health facilities (with clinical observation and interviews with patients and staff).

The findings (both quantitative and qualitative) are presented schematically as sequence of stages at which political patronage and income-seeking are crucial (in the context of a public sector stricken by un-sustainably low wages).

These findings are applicable to ongoing strategic debates: while international donors seek to negotiate the gap between the wealthiest countries and the poorest, their programs entail a further division (between rich and poor, center and periphery) within the recipient state.

In part due to the inequalities that result from intervention, this debate not only has ramifications for Cambodia's new health sector strategy, but also a radical shift in the national bureaucracy, as the country adopts a new reform program that would end the era of government employees receiving salary supplements (of any kind) from international agencies, NGOs or charities.

Keywords: Health; Public Policy; Rent seeking; Patronage systems
Sectoral evaluation: social services evaluation

Local use of national performance indicators: Experiences from the public health sector in Denmark

S. H. Mathiasen, H. S. Pedersen
KREVI, Aarhus, Denmark

The National Indicator Project (NIP) is a performance measurement system, which has been used in the Danish healthcare system since 2004. The system is based on a number of standards and indicators and is meant to further treatment quality within eight health areas. It is mandatory for hospitals to implement NIP although no sanctions are imposed if hospitals do not meet the standards.

This paper reports from an investigation of the organizational consequences of implementing NIP, which revolved around the following questions:
1) How has NIP been implemented on different hospital wards?
2) How has hospital staff experienced this implementation?
3) How can different effects of NIP be explained?

The investigation was carried out as a qualitative case study involving four hospital wards. Data was mainly gathered through semi-structured interviews with staff. The analysis was inspired by Grounded Theory principles and involved the use of ATLAS.Ti.

It shows that NIP is used extensively in two out of the four cases. According to the interviewees NIP has in these cases led to an improvement of the quality of treatment. In the two other cases, however, NIP has not had such an effect.

Several organizational circumstances can explain this difference. This includes management, the attitude towards documentation in general and the extent to which hospital staff concur with the evidence on which the indicators are based.

The study was unable to find significant signs of the negative unintended consequences, which are often reported in the literature on performance measurement.

It is concluded that development of “effective” and “usable” performance measurement systems is indeed possible, but that the success of these systems to a large extent depends on local circumstances in the organizations where the systems are implemented. Thus design of performance measurement systems must in many situations include a focus on organizational development.

Keywords: Performance measurement; Health sector; National Indicator Project; Qualitatative methods

Monitoring and developing the social sector services using indicators – experiences from the social indicator programmes

H. Qvist, P. Rhode, A. Oersted, L. Udsen
Centre for Quality Improvement, The Danish Social Sector Quality Model, 8200 Aarhus N, Denmark

Background
In 2004 the five Danish regions decided to commence development of a number of documentation programmes for the social sector services. The ambition was to gain detailed knowledge about the connection between individual backgrounds, the work provided by the social sector services and the outcomes in regard to the individual's conditions of life.

Methods
The Social Indicator Programmes base their concept of documentation on the particular conditions of social work, by focusing not only on measurable goals but also on activities, context and results. These focal points are clarified by a number of indicators which are based on relevant research, evaluation literature and the involvement of social workers.

To collect data on the selected indicators social workers and the individual service users answer a number of questionnaires. Based upon the quantitative data a number of statistical analyses are done to prepare for professional interpretation and discussion at audit-meetings. Thereby the answers from the questionnaires are interpreted into a social work frame which provides opportunities for learning and quality development and at the same time, assuring the quality of the data collection process and the analyses.

Results
At the present stage six Social Indicator Programmes have been developed covering different target groups. The programmes share a common methodological platform but their content differs considerably. Initial results indicate that the Social Indicator Programmes are not only a useful instrument in deciding best practice in long term; social workers find the programmes a useful communication tool in daily practice, as each programme clearly illustrates an overview of the individual user's problems, resources and development which can be used in planning work and thereby contributes to modifying the service towards the individual user.

Keywords: Indicators; Social sector; Best practice; Knowledge; Evaluation
Finland has – traditionally – been included as Member of the Scandinavian Welfare Societies alongside with Sweden, Denmark, Norway and Iceland. Economic recessions that took place at the beginning of the 1990s as well as during 2008–2009, however, have challenged the foundations of the Finnish welfare model with regard to financial (how much does it cost?), consensual (who supports it?), and functional (what are the strategic choices to be adopted?) aspects.

This paper takes a short retrospective look on the development of the Finnish welfare policies and puts The National Development Programme for Social Welfare and Health Care (Kaste-programme) to the context of policy environment that took place during 1990-2010. The Kaste-programme is the statutory strategic steering tool of the Finnish Ministry of Social Affairs and Health (MSAH) in managing social and health policy issues at the moment. Finnish Government adopted this programme on January 2008. The programme receives its funding from the annual State budget.

The Kaste-programme defines the aims of social and health policy in Finland in 2008–2011. It also defines the focal development and supervision areas, as well as the reforms and legislative projects in social and health care.

The programme seeks 1) to enhance social inclusion and reduce social exclusion, 2) to enhance wellbeing and health, to narrow inequalities in health and wellbeing and 3) to improve the quality, as well as effectiveness and availability of services and to reduce regional inequalities.

The Kaste-programme consists of 39 national measures to support the achievement of the objectives.

The Kaste-programme has been evaluated since 2009. The evaluation approach adopted incorporates various methodological approaches. This paper explores the ways how evaluation has been incorporated into the programme implementation and how the lessons learned by evaluation has been taken into account in programme implementation.
Feeding evaluation results back into policy I

Evaluators and decision makers: some lessons from a self-employment support Programme standard-based evaluation

A. Vergani
Cath. University of Milan, Sociology, Milano, Italy

Background
The background of the paper proposal is double: a) the personal evaluation experience of the Author; b) a formal evaluation conducted in a Northern Italy Region between 2002 and 2006.

Objectives
The aim of the paper is to discuss the relationship(s) between the submission strand subjects with reference to the Public Administrations (PAs) evaluation context in the Italian experience both by a theoretical point of view and considering – as an example – the evaluation of local ESF funded self-employment Programs implemented in Northern Italy between 2002 and 2008. The links between these two focuses, the submission strand and the general theme of the Conference are the following: a) the evaluation relationship(s) that was implemented and occurred with reference to the self-employment support Programs was a multi-actors (or multi-stakeholders) one which was based on the structured interaction between evaluation producers, decision makers (the political level of the local PAs who financed the Programs), users (the technical level of the Pas themselves) and beneficiaries (the self-employment-service supplying agencies and, indirectly, the potential self-employers); b) the Programs’ evaluations (which covered several years and all the Provinces of the Region involved) were conducted on the basis of Evaluation Plans which assumed services’ processes and (final) results standards as a pivotal point.

Results
Main results are: a) the initial definition and on-going revision of evaluation standards were very useful to explicit the single “program theory”; b) regular feed-back of each Program positioning with reference to process and result standards has been a relevant tool for PAs active involvement in the evaluation; c) relationship(s) between Evaluators, Public Administrations, Program implementation agencies and self-employers clearly showed that evaluation cannot be reduced to its technical-methodological dimension but needs to fully assume the role all actors played in the relationship in their heterogeneous profiles and configurations.

Keywords: Standard; Evaluators; Decision-makers; Evaluation plans; Program Theory

Evaluation of science and technology policy mechanism with decision support system: the case of Taiwan

A. L. Cheng
Chung-Hua Institution for Economic Research, Taiwan Economy Division, Taipei, Taiwan

The prevailing framework of logic model for assessing national science and technology development can usually be extended to cover project, program and policy evaluation. However, to put this type of logic framework into practice requires a systematic help with an information system. Especially, when decision makers are to dynamically and responsively react to the changing policy environment, a well-established decision support system is needed. Taiwan is currently reinventing the science and technology development authority and on the way of establishing the DSS for constructing a better-fit national evaluation system. This paper suggests a modified logic framework of evaluation based on the concepts of cost-benefit analysis and endogenous growth theory. In the designed evaluation system, the DSS is seen as some mechanism for providing feedback effects for science and technology policy evaluation and re-formulation. Meanwhile, dimensions of evaluation indicators are constructed to include tangible as well as intangible asset for capacity accumulation. This paper is expected to contribute to add value to academic and pragmatic improvement on overall policy evaluation.

Keywords: National evaluation system; Science and technology development; S&T policy formation mechanism
The EU structural funds programming period of 2007–2013 represents a shift in evaluation practice: From a regulated mid-term evaluation concept, towards a more flexible and demand driven approach called On-going evaluation. In this paper, we would like to share our experiences regarding how Sweden has made its own adaption and interpretation of On-going evaluation, both in theory and in practice, through a model called “Följeforskning”.

The paper aims to address whether or not the Swedish model of “Följeforskning” could be described as a new evaluation concept. And if so, how does the Swedish way of looking at On-going evaluation contribute to the development of evaluation methodology? The questions are answered through a meta-analysis of a large number of evaluation reports on project and program levels, official documents and guidelines, as well as interviews with high-level officials responsible for the development of this model.

It seems clear that “Följeforskning” goes beyond what is stated in the original guidelines for On-going evaluation, drawing on familiar evaluation concepts such as theory-driven evaluation and interactive research, but also beyond typologies like summative and formative. Our findings suggest that this new take on On-going evaluation, suited to specific Swedish conditions and demands, has both benefits and drawbacks: On a theoretical level, “Följeforskning” presents interesting dilemmas about the independence and integrity of the evaluator. Methodologically, “Följeforskning” can be described as the breakthrough of an eclectic, multi-method approach to evaluation. Regarding implementation and use of evaluation results, our analysis shows that “Följeforskning” may facilitate a very quick and direct influence on projects and programmes, drawing yet adding to traditional concepts of participatory evaluation. But this “Swiss knife” type of evaluation also presents a challenge to evaluators as well as to those evaluated, e.g. regarding the definition of roles and responsibilities, insider-outsider dilemmas and the receptory-capacity of those evaluated.

**Keywords:** On-going evaluation; Följeforskning; Swedish evaluation practice; EU structural funds
**How European Governments treat with evaluation processes and results**

**Evaluation and the democratic role of Parliament: Is Parliament in control?**

*B. De Peuter, V. Pattyn, M. Brans*

Public Management Institute, Faculty of Social Sciences K.U. Leuven, Leuven, Belgium

The role of Parliament resides in its democratic control over Government as the commonly generator of policy. While this role contains a general evaluative dimension, Parliament could also use the instrument of policy evaluation in order to perform its key task.

In many countries the controlling power of Parliament is questioned in relation to the executive power of Government. In some countries, Parliament is considered to be active in the advocacy, follow-up and use of evaluation. Parliamentary engagement in evaluation is regarded as a key indicator of an existing evaluation culture within a region or country (Furubo, Rist and Sandahl, 2002).

In this paper we will analyse proceedings of plenary sessions of the regional parliament of Flanders (Belgium) to determine the scope and profile of attention for policy evaluation.

We will firstly examine several dimensions of parliamentary questions: the subject of parliamentary questions on evaluation, the focus of the questions with regard to the steps in a standard evaluation process (structuring, data gathering, data analysis, judgement), whether it focusses more on evaluation processes or on results, whether there is a bias in attention to certain evaluanda. Furthermore we will explore relationships between parliamentary questions on evaluation and the political party affiliation, majority or opposition background. Also we will explore whether the parliamentary focus on evaluation is equally distributed between policy fields.

Secondly, we will screen the answers given on parliamentary questions on possible, planned, going or completed evaluations in order to determine whether there is a match between (governmental) supply and (parliamentary) demand of information on policy evaluations.

**Keywords:** Parliament; Democratic control; Evaluation use; Policy fields

**Shifting purposes of policy evaluation as a key information source for political decision-making? Analysis of trends in evaluation motives**

*B. De Peuter, V. Pattyn, M. Brans*

Public Management Institute, Faculty of Social Sciences K. U. Leuven, Leuven, Belgium

Evaluation literature states a broad array of purposes to evaluate policy measures, projects and programmes. Also different forms of use have been discerned, including accountability and policy learning (e.g. Balthasar, 2006 and Chelimsky, 1997). The literature further identifies institutionalisation of evaluation through structural linkages with the policy process as an important factor for the practice of evaluation (e.g. Jacob and Varone, 2003).

While the role of Parliament is often seen as degrading in power vis-à-vis the Government, the latter could use evaluation to underpin, improve, account for and learn from policy decisions. One key moment in the democratic relation between Government and Parliament is when Government develops and presents its policy notes with upcoming priorities for the new legislature. Policy notes are an opportunity to demonstrate Government’s (formal) engagement in policy evaluation.

In this paper we will analyse all policy notes from three consecutive legislatures (over 15 years) from the regional government of Flanders (Belgium) to determine the volume, types and purposes of intended evaluations. Not only will we compare the profile of explicit evaluation intentions over time, but also a comparison of policy sectors will be undertaken. The analysis aspires to identify shifts and trends in evaluation purposes over time and reveal similarities and differences between policy sectors. Further, explanatory factors will be explored such as the NPM-inspired government wide reform of ‘better administrative policy’ which has put evaluation more explicitly on the list of key competences of the sectoral departments. On the one hand the authors want to examine whether an increased attention to evaluation can indeed be observed in practice. On the other hand it will be analysed whether the administrative reform has introduced a shift in the relative weight of different evaluation purposes.

**Keywords:** Evaluation purposes; Planning; Accountability; Policy learning; Institutionalisation
Evaluating policies, understanding mechanisms and cumulating evidence

Brad Astbury1, Sandra Speer2, Gill Westhorp3, E.L. Leeuw (Chair)4

1University of Melbourne, Melbourne Graduate School of Education, Melbourne, Australia
2Independent Evaluator, Wiesbaden, Germany
3Community Matters, Mt Torrens, Australia
4Maastricht University & National Institute of Justice Research, Den Haag, Netherlands

Within this symposium Brad Astbury will present a paper on the role of demi-regularities and Context-Mechanism-Outcome Pattern Configurations (CMOCs) in evaluation. Realists prioritise their evaluative efforts on identifying underlying ‘mechanisms’ which in particular ‘contexts’ generate ‘outcomes’. These four CMOC linked concepts form the ‘signature argument’ or analytic centrepiece of realistic evaluation and realist synthesis (Pawson, 2006). A particular focus will be on the idea of portable mechanisms and their utility in enhancing the transferability of social interventions. Sandra Speer will give an overview of CMO logic from the perspective of economics. In economics, CMOs are not usually discussed explicitly; however, they are implicitly often the basis of economic thinking. Secondly, also economic approaches for evaluation exist, which follow a kind of CMO logic. Gill Westhorp will focus on human services programs (those in education, health, welfare and so on), which often intend to create change at multiple levels of at least two systems – the service provision system, and the client system. The mechanisms of change at each level of the system are usually different. This paper will use one example to outline a model for considering CMOCs at different levels of systems and the interactions between them.

This symposium shall bring together different perspectives on CMOCs and their importance for opening up “black boxes” in a variety of evaluation domains. The presenters will also discuss the practical implications of this work for contemporary evaluation practice.

Keywords: CMO logic; Realistic evaluation; Demi-regularities; Economics
Evaluation policy: Regional development I

How an integrated approach to evaluation and communications can contribute to national and regional development

G. Ellis, M. Pai
Gellis Communications, Strategic communications, 1050 Brussels, Belgium

URBACT, the EU programme for the promotion of dynamic, sustainable and cohesive cities in Europe, financially supports projects that aim to make European cities better places to live. In September 2007, Gellis was hired by the URBACT Secretariat to evaluate its communications between 2003 and 2006 and provide strategic advice.

During the period of data collection, it was discovered that unearthed problems were a serious threat to the visibility and the reputation of URBACT, not only by limiting its appeal amongst potential beneficiaries but also for leading sponsor organisations to question the very relevance of the programme. These problems included the lack of a strategic approach to communications, incorrect identification of target audience, and confusion amongst the URBACT Secretariat’s staff between the notions of knowledge management and communication.

Using the Logical Framework Approach, Gellis redefined the goal and objectives of URBACT’s communications: to enhance the reputation of URBACT, thereby contributing to the achievement of the programme’s objectives. Considering the structural problem of URBACT’s communications as well as the limited human resources of the Secretariat, a very gradual strategy was adopted to fix issues by order of priority:

Priority 1 – Strengthening internal capacity
Priority 2 – Developing multimedia content
Priority 3 – Empowering Beneficiaries

This strategy is proving successful, with the most striking results including:

- A sharp increase in the number of media articles about URBACT, from 10 in 2008 to 70 in 2009.
- Higher website traffic figures (from 310,000 single visits in 2008 to 360,000 in 2009);
- Average number of conference participants jumped from 300 in 2008 to 450 in 2009

Keywords: Urban development; Strategic communications; Target group classification; Internal capacity; Multimedia content

Regional development policies and performance reserve system: searching for evaluation elements

F. Mazzeo Rinaldi, A. Spano, G. Giorgio
1 University of Catania, Sardinia Regional Evaluation Unit, Catania, Italy
2 University of Cagliari, Sardinia Regional Evaluation Unit, Cagliari, Italy
3 University of Sassari, Sardinia Regional Evaluation Unit, Sassari, Italy

In 2007–2013 the Italian regional development policy focused on improving basic services in four strategic areas such as Education, Child and Elderly Care, Waste management and Integrated Water services, according to the EU performance reserve mechanism. Three billion euros have been allocated to be awarded as monetary prizes to the eight Mezzogiorno regions that will achieve the set objectives and targets in those service areas. The path to achieve the four objectives is assessed through eleven statistical indicators, which measure the quality and quantity of services provided to citizens. Compared to the previous performance reserve period (2000–2006), the main difference is related to indicators and targets bonded not longer to improve capacity-building in public administration, but to service improvement.

A first assessment was made in 2009 (based on 2008 data), as requested by the performance reserve system, for the allocation of a first portion of the prize, for a total of 640 € million. Based on this first allocation, the objective of this paper is to assess the implementation level of Action Plans (AP) developed by the eight regions for the objective “Child and Elderly Care Services”, on the basis of the associated statistical indicators.

To achieve this objective, a comparative analysis of the AP implementation process will be carried out, together with the performance indicators achieved. A cluster analysis will help to highlight common elements in the behavior of the organizational units responsible for citizen service delivery.

The paper allows a first comparative assessment of the implementation level of APs developed by the eight regions in the south of Italy, in the framework of performance reserve objectives, as well as the contextualization of performance indicators through an exploratory data analysis which aims at sorting the different organizational units into groups.

Keywords: Performance indicators; Performance reserve mechanism; Cluster analysis; Child and elderly care services
Regional level, ex-ante social impact evaluation studies and methodologies

J. M. de Vet¹, K. Berden², S. Tamminen²

¹ECORYS Brussels, Director Brussels, Belgium
²ECORYS Netherlands, Macro and Sector Policies, Rotterdam, Netherlands

Ex-ante impact assessments/evaluations are an increasingly important tool for making better policies at EU, national and regional level. Over the past few years, significant efforts have been made to improve the quality of these impact evaluations, especially at EU level, and, at least in theoretical level, there are various quantitative methodologies that can be applied.

Within our recent work for the European Commission, we have 1) made an overview of regional level, ex-ante social impact evaluations and methods, 2) created a structured “Roadmap” to guide systematic analysis of especially regional employment and redistributive impacts and 3) within the context of this “Roadmap”, provided an overview of relevant methods for these evaluations. This roadmap is presented below.

We found that in Europe the methods used for and the quality of social impact evaluations at regional level have varied. Particularly the use of quantitative methods has been limited in some countries, while very sophisticated methods have been developed. At Member State level, the political interest in and the demand for social impact evaluations has been comparatively weak combined with frequent tight time and budget limits.

For the methodology selection a good understanding of the causal links and their strengths is important. While quantitative methods can provide valuable estimations on the size of the impacts, modelling approach should only be considered if causal relations are clear, impacts are likely to be strong and needed data available. Further, most of the current models are in national level and model builders and users have pointed to significant hurdles on up-scaling the models to be applicable in whole EU level.

The overall use of ex-ante social impact evaluations and the findings of our studies will be discussed in a debate by the authors and a representative of the authorising client, European Commission, DG Employment.
Participatory evaluation

Participatory assessment of development – a new impact evaluation methodology for international development that overcomes some common pitfalls of evaluations

W. Rijneveld¹, D. De Groot¹, F. Zaal³, T. Dietz⁴
¹Woord en Daad, Advice and Research, Gorinchem, Netherlands
²ICCO, Research and Development, Utrecht, Netherlands
³KIT, Sustainable Economic Development, Amsterdam, Netherlands
⁴University of Amsterdam, Amsterdam Institute for Metropolitan and International Development Studies, Amsterdam, Netherlands

Background
Impact evaluation is usually linked to specific interventions. Impact to be measured is often predetermined by those who initiated the intervention. In international development with its inherently unequal power relations this raises the political question who defines impact. Secondly, it raises the methodological dilemma of choosing between objective, externally defined operational variables for development often far away from beneficiaries, or using participatory methods that easily suffer from bias, because of the dependency relation. Thirdly, the focus on a single actor chain easily leads to a tunnel vision where the object of research is insufficiently placed in a complex relation to other local or international interventions and contextual shocks and trends. This diminishes the strategic usefulness of conclusions.

Methodology
A methodology is being developed that takes the local community as point of departure. Shocks and trends over the past generation are assessed by means of participatory workshops. Complete lists of interventions are developed and each intervention is assessed. The gendered assessment includes the diversified impacts on different locally defined wealth groups, impacts on different livelihood capitals, shifts in impact over time and the relations between contextual changes and interventions.

Results
The methodology has been developed in nine workshops in Ghana and Burkina Faso. Analysis of all data provides a wealth of information about local perceptions of impact of interventions of different actors in relation to contextual changes. The paper contains a description of the methodology, some illustrative findings, a discussion of the methodological strengths and weaknesses and a positioning in the field of evaluation methodologies.

Application
This impact evaluation methodology adds value in the range of methodologies in overcoming some pitfalls of existing methodologies and bringing the policy discussion about development back to the heart of the matter: what is development and who is in charge of it?

Keywords: Impact assessment; Participation; Contextual change; Livelihood capitals; Actor types
A multi-level participatory evaluation: experiences from a developing country – Philippines

O. Ocleasa
Acdivoca, Monitoring & Evaluation, Davao City, Philippines

This paper presents experiences and lessons learned in Participatory Monitoring and Evaluation in the Philippine setting.

One is the Multi Level Participatory Evaluation, an innovation from the traditional midterm review which uses external experts to assess the project status. Instead the stakeholders, project team and the beneficiaries reviewed and assessed the project at different levels particularly its outcomes and impacts or emerging impacts. This is composed of several processes or activities as the Stakeholders Participatory Assessments Workshops for stakeholders to come together to jointly assess the impacts, the Project team and field implementers workshops, the top level management joint review, the partners level and the village-based beneficiaries participatory impact assessment. All their outputs were consolidated into one report which is now the Participatory Midterm Evaluation Review report. Same processes were conducted at the end of project period. The results were assimilated into the End of project report made by Independent/External consultants. However this method still included the review of project documents and reports generated throughout the implementation. An actual case of this was done by the presenter on their project the Sustainable CoCoa Enterprise Solutions for Small Holders (SUCCESS) Alliance Philippines a USDA funded project implemented by ACDI VOCA an international NGO based in Washington D.C., USA with its Philippines office. SUCCESS Alliance is a project which is national in scope.

The paper will also share the experiences of the presenter in Participatory Impact Monitoring in the previous project funded by International Fund for Agricultural Development, the Western Mindanao Community Initiatives project (WMCIP) with the Department of Agrarian Reforms (DAR) of the Philippines Government which was implemented in conflict-affected areas in Mindanao. Lastly it will present the Participatory Farmers Adoption & Results Monitoring (PARFARM) which is one of the 9 winners of the 2010 ALINE Farmers Voice Awards given by the Agricultural Learning & Impact Network (ALINE) a project of Institute for Development Studies (UK) and Keystone Accountability with funding from Bill & Melinda Gates Foundation.

Please check this:
- http://www.aline.org.uk/awards
- www.acdivoca.org
- www.thesuccessalliance.org
- www.ifad.org

Keywords: Participatory evaluation

Challenges of Measuring Impact Evaluation of Community Based Participatory Research Programs: A Case of Bukoba Rural, Republic of Tanzania.

N. Singh, D. Ettling
University of the Incarnate Word, Nutrition, San Antonio, USA

Background
Impact evaluation is intended to determine more broadly whether the program had the desired effects on individuals, households, and institutions and whether those effects are attributable to the program intervention. Of particular interest is the impact that community-based programs are having on rural women's economic and social empowerment in the Bukoba rural, Kagera Region, Republic of Tanzania. The gender dimension of poverty in the United Republic of Tanzania is striking. 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) is very limited, which increases their income and food insecurity.

Methods
Reach-Out Africa was initiated in Tanzania in 2003 with a primary goal of women's empowerment through capacity building to effect economic sustainability and social empowerment. Emphasis has been placed on building women's business and leadership skills with the intention to improve their health and economic status and impact their role in family and community.

Objectives
This paper addresses the challenge of measuring the actual social impact on rural women of capacity building activities delivered over a five-year period (2004–2008).

Results
We have created a review-based list of impact indicators that we believe more accurately depicts the positive changes resulting from capacity building projects of such a nature. As an example, the following are some of the indicators that we have found useful in determining impact in rural Bukoba: a) involvement in major family decision making; b) Participation in public action; c) Physical Mobility/Social Interaction; d) Political and legal awareness; and e) Ability to make small and large purchases.

Application
There is a great need to continue the evolution of useful and recognized tools to measure the impact of social development. This forum offers the opportunity to discuss this particular case in Bukoba, Tanzania and further understanding in the field.

Keywords: Community-based Projects
Evaluating Impacts of a Malaria Prevention Program (Long-Lasting Insecticidal Nets (LLIN)) in Africa: Methodological Innovations Involving a Partnership/Collaborative Approach

H. Cummings¹, H. Paulmer², D. Henderson³

¹University of Guelph, Environmental Design and Rural Development, Guelph, Canada
²Harry Cummings and Associates, Senior Evaluator, Ottawa, Canada
³Canadian Red Cross, International Programs, Ottawa, Canada

Background
Malaria prevention is a major requirement of many countries in Africa. This paper presents the results on an evaluation of the impact of the LLIN distribution programs carried out by the Canadian Red Cross and its partners between 2005 and 2009 in Mali, Madagascar, Nigeria, Sierra Leone and Togo supported by CIDA. The evaluation focus was on the capacity building impact of the projects, and the impact of the projects at the household, community and district level. The paper also focuses on methods used.

Objectives:
the research; demonstrates the impact of the program, shows what methods were used and suggests their utility for future evaluations, and demonstrates how mixed methods evaluations work in impact assessments.

Methods:
The methods used include; participatory collaborative workshops, focus groups, population surveys, key informant interviews and stakeholder surveys. The focus is on the elements of a successful impact evaluation.

Results:
A collaborative evaluation workshop, held over six days in Mali, proved to be an effective way to assess the impacts of the project. Having all countries, external evaluators, the CRC, Ministries of Health, and the National Red Cross societies together for an intensive and extended period of time allowed issues to be explored in an open and honest way. National societies capacities were developed in all countries through the campaigns.

Conclusions:
The results were easy to document because of; the completion of the project in all countries in advance of the workshop, the completion of large baseline and post campaign surveys in all countries, the low likelihood of the CRC doing similar projects in the same countries in the near future, the combination of quantitative and qualitative methods used in the evaluations and the triangulation associated with mixed methods, the presence of outside facilitators to elicit fresh and frank opinion and fact, and other factors

Keywords: Impact; Malaria; Africa; Collaborative evaluation; Mixed methods
Evaluation of development as social change

Quality evaluation of development as complex social change: rethinking rigor based on values, standards, tradeoffs and consequences

I. Gujit1, Z. Ofir2, R. Rutten3, R. Davies4

1 Learning by Design, Randwijk, Netherlands
2 Independent, Geneva, Switzerland
3 Cordaid, Policy, The Hague, Netherlands
4 Independent, Cambridge, United Kingdom

Many efforts in development are innovations, rather than the rolling out of guaranteed successes. This means they are more complex – i.e. emergent and unpredictable – and therefore inherently risky, yet critical if development is to be sustained and respond to urgent issues that require innovative responses. Current evaluation practice has not yet embraced the full implications of assessing ‘the complex’. Existing approaches fall woefully short and growing attention on evaluative efforts that promote a reductionist notion of development are worrying.

How we view processes of social change and development and what we consider quality evaluation of interventions to achieve these are inextricably linked. There is a risk that efforts to improve the quality of development evaluations will only support the evaluation of standardized, simple interventions. Solid debate is needed to find better ways of evaluating ‘the complex’ without compromising quality standards, including understanding better the impact of methodological choices on the processes of social change.

Four panelists will take on the question of how rigour of evaluative practice can be rethought, given complexity of social change. Dr. Z. Ofir will discuss the challenge of maintaining space for diversity of impact evaluation approaches. Dr. I. Gujit will share key insights from an international conference held in May 2010 that brought together over 150 experts. The methodological options showcased and debated offer rich insights for others keen to enhance the quality, including rigour, of evaluation, while respecting the nature of complex social change. Ms. R. Rutten will share Cordaid/CIDIN’s experiences with participatory impact assessment in health and education, using purposive sampled control trials in three countries. The trade-offs, real time and context-specific methodological choices will be discussed, in relation to what insights about the impact of development were made possible. Dr. R. Davies will act as discussant for the session.

Keywords: Complexity; Social change; Rigour; Methodological choice; Case studies
How to capture effects in real world public policies

Capturing the Effects of Europe’s Cohesion Policies

M. Hart, R. Picciotto, V. Gaffez, M. Saunders

According to Pawel Samecki, EC Commissioner for Regional Policy, “a strong Union needs policies that ensure that everyone can benefit from integration”. This imperative underlies the resurgence of interest in the effects of cohesion policies and programs in the expanded European space. While mixed methods of increasing sophistication have become available, commissioners of structural funds evaluations continue to express frustration with the lack of sound and credible evidence of practical use to policy makers.

What are the effects of EU programmes that span infrastructure development, enterprise support, human resource development, urban regeneration, etc.? Are these programs responsive to cohesion policy goals and citizens’ felt needs? Do they deliver value for money? Do they translate into concrete, measurable effects that can help satisfy public expectations? What are the implications for EU evaluation methods and practices?

To help trigger an open debate on the dilemmas associated with public officials’ demands for simple, accurate and rigorous measures that capture the effects of EU cohesion policies, a panel chaired by Murray Saunders former President of the European Evaluation Society will present three contrasting perspectives. First, Ms Veronica Gaffez, Head of Evaluation in DG Regional Policy and Acting Director, Policy Development will set the stage by providing a bird’s eye view of the challenges she faces on a day to day basis in overcoming the obstacles to evaluation utilization in the real world of EU policy making.

Next, Professor Mark Hart, Professor of Small Business and Entrepreneurship in the Economics & Strategy Group of the Aston Business School (UK) will challenge the prevailing orthodoxy and highlight the risks of evaluation methodologies that pretend to attribute measured outcomes to complex and evolving social programs. Finally, Robert Picciotto, EES board member and former Director General of the World Bank Independent Evaluation Group, will identify hard-won development evaluation lessons that may be relevant to the evaluation of EU’s social, regional and cohesion funds.
Tools and methodologies II

To Whom are We Accountable and for What? Evaluation of Non-profit Organizations and Local Wellbeing

K. Piirainen, V. Kuvaja-Köllner, A. Kettunen
Diaconia University of Applied Sciences, Research Centre for Social Economics, Pieksämäki, Finland

In Finland, traditionally the relationship between the public sector and non-profit organisations in the area of social and health care have been strong. The responsibility of municipalities is to arrange social- and health care services, but they use the possibilities of non-profit organisations to support citizens wellbeing. The decline of public economy has raised the question of the accountability of non-profit organisations and thus, the needs to develop methods and models to evaluate the effects and costs of non-profit organizations. When considered from the organizations point of view, municipalities can not dictate the targets and the points of emphasis of these organizations. Non-profit organizations posses an independent position in the core of civic society.

The objective of the presentation is to help understand the challenges in accountability evaluation and methods used for practical purposes when the object of evaluation is the non-profit organization.

This study described a model for evaluating the effects and costs of non-profit organizations of social and health care. The model was created in the development project through pilot evaluation. The aim of the project was to develop local organisation’s own facilities to show the worth of their actions in creating wellbeing.

The character of the accountability evaluation was determined mostly by the organizations’ targets. We concluded to classify the targets in three ways, i.e., the benefit of the society, the wellbeing of the organization’s own members or supporting the groups that are under the risk of exclusion. Evaluation methods and possibilities depend decisively on the focus of organizations and their relationships to their stakeholders. The challenges for the accountability evaluation in the future should be discussed in general but especially in terms of the non-profit organisations.

Keywords: Evaluation; Accountability; Non-profit organizations

Experience with systems constellations in theory-based evaluation – a new tool to support actor-centred evaluation

U. Kopp, A. Martinuzzi
Research Institute for Managing Sustainability WU Vienna, RIMAS at WU Vienna, Vienna, Austria

Theory-based evaluation has become a widely used method for programme evaluation: programme theories, logic models and frameworks are used to analyze how programme activities lead to outcomes and impacts, taking a rather passive view on the programme theory. Although programme theories are often developed through participatory processes, in most cases the actors involved in the programme are not being considered in detail. In order to get a holistic view of a programme, to understand all its mechanisms, and to be able to analyse the implicit and explicit programme theory, it is essential to reconstruct the orientation and motivation of the actors involved. The tool of systems constellations, already well established in systemic consulting and organisational development, enables the evaluator to identify, visualise and consider the actors’ perspectives on the programme in a workshop setting with the participation of those responsible for the programme’s success. In the framework of ongoing research funded by the Austrian National Bank we studied the potential of the usage of systems constellations in theory-based evaluation as well as with methods as varied as constellation analysis, rich pictures, soft systems methodology and multi-criteria analysis in several evaluation cases. Our contribution is going to present results reached so far, indicating inter alia that this tool is useful especially as a heuristic tool to rapidly give an insight into formal and informal relations between programme actors and providing an opportunity to reveal and to visualise their implicit knowledge. The insights gained during constellations work offer a basis for more actor-centred programme theories and evaluations.

Keywords: Theory-based evaluation; Systems constellation; Actors-perspective; Systemic; Participatory evaluation
Matching Systems Methods To Evaluation Questions

R. Hummelbrunner
OEAR Regionalberatung

Background
Interest in systems approaches has been growing in recent years, but application often lags behind this interest. Part of the problem is that the range of systems methods is so large that it’s difficult to select methods that are appropriate to a particular situation. The presenter is co-author of a new book (with Bob Williams) – to be published in autumn 2010 – “Systems Concepts in Action: A Practitioner’s Toolkit”. This book addresses this problem directly and proposes evaluation questions as a solution.

Objectives
The paper presents a tool that allows participants to choose appropriate systems methods to address specific evaluation questions and which is based on the structure of the above mentioned book. Systems methods are grouped around types of evaluation questions which and thus shows how evaluators can match specific systems methods to their evaluation questions.

Methods
The paper will start with an overview of systems approaches and the circumstances in which different methods are applied. It will then move to a more detailed consideration of the kinds of evaluation questions that particular systems methods address, by grouping systems methods around types of evaluation questions. The presenter will also provide additional advice and insights on particular systems methods from his own practice.

Results
As a result of this paper participants will be given a tool for identifying appropriate systems methods on the basis of evaluation questions.

Conclusion/Application to practice
The paper gives guidance on which systems method would be suitable in which evaluation situation. Thus it not only contributes to the problem of entwining evaluation and systems approaches, but also highlights a trend in both the evaluation and systems field: to bring multiple methods and multiple methodologies to bear on an individual inquiry. This question oriented method of selection promotes this approach.

Keywords: Systems methods; Evaluation questions; Multiple methods
Evaluation and governance I

Why is it so hard to make evaluation a regular aspect of good management?

M. Laeubli Loud, P. Zeyen, E. Bruhin
Swiss Federal Office of Public Health, Evaluation and Research Management, Liebefeld, Switzerland
marlene.laeubli@bag.admin.ch

Issue/problem

With the growing use of results based management in the public sector coupled with the development of performance audit and strategic performance review, there has been increasing concern about the implications of such changes for the evaluation profession. What is the added value of evaluation? What is its complementary role? How indeed should it position itself for the future?

This paper will argue for the need to transform evaluation's traditional roles and find an approach which can provide a healthy balance between facilitating organisational learning and building evaluation capacity, on the one hand, and independent verification on the other.

A possible solution – the hybrid evaluation approach

Mark Bittel's “hybrid evaluation approach” promises to do just this. It combines the best of both the internal (self) and external evaluation worlds, but requires a role change for the external evaluator – from “doing” to “supporting” the evaluation process. The hybrid evaluation process encourages organisations to develop the use of evaluation logic and skills to foster and enhance practice. It builds reflection into practice in order to create a learning organisation, ensuring that evaluation is done by organisations themselves rather than being done to them (by external evaluation organisations). In turn, this leads to increased ownership of findings and puts the responsibility to act firmly back into the hands of delivery organisations. It is mainly about contracting in the thinking and learning, rather than always contracting it out – this being its greatest and most useful attribute.

However, Mark Bittel’s approach was developed and applied in the UK context. Our paper discusses its transferability to a federalist context and reports on progress after two years of experimentation.

The presentation will end with a synopsis of the learning from the process, acting as a catalyst for wider discussion.

Keywords: Evaluative-thinking; Hybrid evaluation; Health policy

Promoting evaluation culture in organisations; roles for evaluators

J. Owen
The University of Melbourne, Centre for Program Evaluation, Melbourne, Australia

Theoretical issue.

A key objective for many evaluators is to enhance the commissioning and use of evaluation findings within organisations. Despite the interest in evaluative capacity building over the past 10 years, there are still unanswered questions about how an organisation adopts an evaluative perspective within its operations. How does the rhetoric become organisation practice?

Methods.

This paper relies on experiences related to factors that contribute to organisational evaluative capacity and the creation of an evaluation culture. Using research from work with three organisations over several years, the study used inductive methods based on three cases to develop findings about organisational take-up of evaluation over time.

Conclusions and Applications.

A promising concept on which to base the analysis was that of sustained interactivity, the means by which professional evaluators work over time with organisational leaders to embed the routine use of evaluation findings in organisational decision-making. These processes require evaluators to have a wide range of interpersonal skills, not usually associated with traditional applied social research.

The presentation will discuss how these processes have been used to integrate evaluation into decision-making about (i) policies and programs that are the subject of a given evaluation, and (ii) lessons from these evaluations for organisational structures and ways of operating that affect the delivery of future or new interventions that an organisation has in mind.

Keywords: Organisational Evaluation Culture Promotion; Evaluator Roles
Learning to use evaluation: the role of the EU and of regional institutions

*M. Muru*

University of Bristol, Department of Politics, Bristol, United Kingdom

This paper is an ongoing research on the development of evaluation capacity in the EU regions receiving support under the EU Structural Funds. In some of these, EU cohesion policy has been the sole driver of evaluation and it seems that after more than twenty years from its introduction the evaluation capacity of some regions is still weak. A possible reason might be that these regions do not receive effective and incisive guidance from the European Commission, despite their surely commendable effort to promote the practice of evaluation across Europe. Another reason could be that at sub-national level there is still some resistance to the use of empirical evidence to support the policy-making process. This analysis adopts the multilevel governance standpoint, according to which there is a continuous interaction among the various level of governments taking part in the design and the implementation of cohesion policy. The first aspect examined is how the Commission elaborated their evaluation requirements and guidelines. The analysis covers in particular the epistemological position of these and the relation with the EU evaluation stakeholders (other EU institutions and the net contributors to the EU budget in particular). The second aspect examined is how the EU evaluation requirements are implemented in the EU regions, how these offer guidance and how the communication between regions and the Commission takes place in the specific case of the evaluation requirements. Furthermore, this paper will examine how the evaluation requirements affect the regional policy-making process and the role of local stakeholders in influencing both the policy design and outcomes. The result is that the Commission offers an indispensable input and also influence the regional policy-making process. However, although there have been significant improvements, some constraints at regional level hampers the full development of evaluation capacity.

**Keywords:** Cohesion Policy; Regional Evaluation Capacity; Multi-level governance

Evaluating network governance

*E. Vedung, N. Hertting*

Uppsala University, IBF, Institute for Housing and Urban Research (Gävle) and Dept of Government (Uppsala), Sweden

Network governance (joined-up government, co-production) is a crucial management idea in our time. Perfect network governance is characterized by self-organization, continuous horizontal cooperation among equals and the absence of a central governor cum coordinator. Complete network management occurs when a set of interdependent actors voluntarily, informally and trustingly work together around substantive problems in order to pursue their own and their common interests. Ideal-type network governance involves a number of actors in a flat organization without a common overriding center who by communicating in a frank and honest manner probes opportunities to voluntarily coordinate themselves by merging or exchanging resources with each other to solve problems.

Perfect network governance has four defining features:

1. With reference to substantive and institutional complexity surrounding social problems policymakers regard coordination across program, agency, sectoral and organizational boundaries as an important purpose of governance.
2. Since horizontal, self-organized cooperation among directly involved stakeholders is regarded a basic mechanism for this coordination, the operational actors themselves are assumed to have incentives initiate and maintain cooperation for their own and for the common good.
3. Since network management operates across formal hierarchies and agency jurisdictions in government and between the public and the private spheres, open, flat and informal forms of coordination are promoted as appropriate, with maximum opportunity to access (entry) and defection (exit).
4. In order to launch, facilitate and maintain such horizontal forms of cooperation key actors may indirectly stimulate through any type of instrument.

How should network governance be evaluated? Should it be evaluated by its results or by its processes? What value criteria should be used? These and other evaluative issues will be discussed in the paper.

**Keywords:** Network governance; Network management; Joined-up government; Co-production; Horizontal coordination
This symposium presents the results of three synthetic reviews, funded by the International Initiative for Impact Evaluation (3ie). 3ie Synthetic Reviews examine the range of existing evidence on social and economic development interventions conducted in low- or middle-income countries. They are conducted to Campbell Collaboration standards of systematic review, developed from a published study protocol and employ a theory-based approach to assessing programme effectiveness.

-- The effects of microcredit on women’s control over household spending in developing countries: a synthetic review study

Over the last two decades programs offering microcredit and associated financial services have spread across the world, reaching millions of people. One of the questions receiving prominent attention in the literature has been whether microcredit can empower women and change existing gender relations. Combining Campbell standards of systematic review with a realist synthesis approach, we analyze in what circumstances microcredit can make a difference regarding women’s decision-making power over household expenditures.

-- Effectiveness and relevance of interventions designed to reduce the prevalence of female genital mutilation/cutting

Female genital mutilation/cutting (FGM/C) is a traditional practice that involves excision or other injury to the external female genital organs. Using an integrative synthesis evidence approach (Figure 1), we conducted a systematic review of the literature which assessed the effectiveness of interventions designed to combat FGM/C, and the extent to which intervention programmes had heeded and built upon factors related to the continuance and discontinuance of FGM/C.

-- The impact of agricultural extension services

Agricultural extension and advisory services are an important element of efforts to improve quality of life in developing countries. This paper synthesizes quantitative and qualitative studies examining the circumstances in which these programmes foster improved farmer outcomes and why. Evidence is synthesized on effectiveness from rigorous impact evaluations, and on the factors underlying effectiveness of farmer field schools.

Keywords: Impact Evaluation; Systematic Review; Microfinance; Female Genital Mutilation/Cutting; Agricultural Extension
**Sectoral policy area: Health**

**Evaluation and Health Technology Assessment – Potential for mutual learning?**

**K. Vrangbaek, N. Permin Berger**

AKF Danish Institute of Governmental research, Health, Copenhagen, Denmark

Health Technology Assessment (HTA) deals with the evaluation of new and existing health technologies, with the aim of providing input into decision processes. It is an explicitly applied type of evaluation research (Banta and Luce, 1993). Recently, HTA perspectives have broadened from a focus on “hard” technologies, such as pharmaceuticals and medical devices (Reuzel and Van Der Wilt, 2000), to cover “soft” or “social” technologies, such as the organization and sequencing of interventions. There has also been a concurrent re-orientation in the purpose of HTA, shifting from ‘technology-driven’ to ‘problem-driven’ inquiry. Instead of focusing on a single technology for a given problem, a range of technologies is considered and compared (Kristensen and Sigmund, 2008).

Given this expansion and re-orientation, the Danish approach has been to include organizational and patient perspectives into the analysis, seeking a broad understanding of the use and impact of technologies beyond the well-established clinical and health economic methodologies, leading to a sometimes uneasy marriage. Research is needed into making the different parts of the HTA work together; synthesizing and weighing the results from the four analyses resting on different ontological and epistemological underpinnings remains a challenge.

The evaluation field has undergone developments that mirror the evolution of HTA. Evaluation moved from a narrow, technical focus to include a broader range of questions, drawing on new sources of evidence and thus new methods and standards of knowledge production, synthesis, generalization, and utilization (Guba and Lincoln, 1989, Shadish et al 1991, Alkin and Christie 2004, Donaldson, Christie, and Mark 2009). This paper discusses whether evaluation literature can address some of the key challenges for HTA. The specific research question is thus: What are the main methodological and policy dilemmas as HTA develops beyond clinical and economic evaluation strategies, and can evaluation theory assist in responding to such dilemmas?

**Key words:** Health Technology Assessment; Evaluation Theory

---

**The need for coordination between primary and secondary health care – An evaluation of follow-up home visits to elderly persons**

**U. Hjelmar**

AKF Danish Institute of Governmental research, Copenhagen, Denmark

Studies have shown there is a growing need for coordination between primary and secondary health care. Follow-up visits to elderly persons are an example of a programme trying to address this issue. The aim of follow-up visits by the general practitioner and district nurse (within a week after discharge from the hospital) is to reduce hospital readmissions and improve the overall wellbeing of the patient, and it has been shown that there is a strong evidence of effectiveness in these programmes. At the same time, studies have shown that it has been difficult to implement these programmes because of a number of organizational obstacles, including coordination between the organizations involved in the process.

The aim of this paper is to examine coordination issues between primary and secondary health care on the basis of a particular case, a programme in Copenhagen, Denmark, trying to implement follow-up home visits to elderly persons. Focus in the programme evaluation is on coordination processes at the operational level and the actors’ interpretation of these processes. The analysis is based on in-depth interviews with hospital staff, general practitioners, and district nurses. Theoretically, the evaluation is based interorganisational network theory trying to explain how organizations can work together and, more specifically, conditions of network formation between organizational systems.

The results show that there is a lack of integration and poor communication between primary and secondary health care. This is mainly due to problems with the formal instruments securing coordination between hospitals, general practitioners and district nurses. At the same time different working cultures and motivation structures in the different organizations make interorganisational coordination difficult. These interorganisational factors need to be addressed in future programmes in the area in order to fully exploit the potential health benefits.

**Key words:** Health policy; Programme evaluation; Interorganisational coordination; Integrated care; Home visits
Measuring impacts of innovative and emergent social policy programmes: lessons from evaluating the English “Skilled for Health” programme

K. Junge, D. Hills
Tavistock Institute, Evaluation and Research, London, United Kingdom

For the European Commission and EU member states tackling health inequalities is an important public policy issue. In England, the Department of Health and Department for Innovation, Universities and Skills jointly ran the second phase an innovative and emerging programme – Skilled for Health – aimed at reducing health inequalities by improving health literacy amongst the most disadvantaged people in England and engaging those not traditionally involved in adult learning.

Drawing on the experience of the evaluation of this programme, this presentation explores the challenge of generating robust outcome data from a complex and complicated interventions that have broadly defined objectives and display a high degree of diversity in implementation.

The external evaluation of Skilled for Health was charged with assessing the impact of the intervention at programme level and measuring outcomes on individuals. The evaluation design accommodated the diversity of the programme (in terms of learners, course formats and locations) by using a theory of change framework and a co-production approach to the core component of the evaluation: a measurement tool to assess individual and programme level outcomes. This evaluation design resulted in a context sensitive understanding of the distance travelled by the programme beneficiaries and a much tighter definition of the Skilled for Health intervention, its key outcomes and how they come about.

These evaluation results support a possible future randomised controlled trial of the Skilled for Health intervention which would deliver the hard data on the basis of which further investments can be justified. The implications for those designing and funding innovative and emergent programmes on health inequalities and other social policy challenges are that a pilot phase and mixed method evaluation designs are promising strategies to help intervention to “settle down” and create robust outcome data.

Keywords: Health inequalities; Theory of change; Impact evaluation; Complicated and complex interventions

Social insertion and habitation: a way to the evaluation of housing situation for people with severe mental disorders in Brazil

J. P. Furtado, A. Tugny, C. Genoroso, A. Maris Guerra
1 Federal University of Sao Paulo, Sao Paulo, Brazil
2 Universidade Federal de Minas Gerais – UFMG, Escola de Belas Artes, Belo Horizonte, Brazil
3 Universidade Federal de Minas Gerais, Psicologia, Belo Horizonte, Brazil

In what it concerns to the delivery of housing needs, people with severe mental disorders in Brazil have two possibilities: the Therapeutical Residential Services (SRTs), or alternative forms from family support and/or from the mental health services, such as the Psychosocial Support Centers (CAPS). It seems fundamental evaluating how people with severe mental disorders constitute their habitation (habitus) and their social insertion in these experiences, standing from structural elements of housing (shelter, privacy, safeness and comfort) and support (social network and health services), as well as the treatment delivered by health services, in a de-institutionalization context. To make it possible, there will be presented and discussed some principles, analysis categories and possible contributions for a method of qualitative and interdisciplinary evaluative investigation developed by researchers from different knowledge areas (anthropology, architecture, psychoanalysis and public health), by providing a more systematic approach on this important thematic.

Keywords: Supported housing; Interdisciplinarity; Social Insertion; Psychiatric Reform; Housing
**Evaluation policy: Regional development II**

Evaluating the contribution of environmental investments to regional development. The case of the European Regional Development Fund over 2000–2006

*M. Van Overbeke*

ADE S.A., Louvain-La-Neuve, Belgium

The paper's main topic is the role the ERD-supported environmental investments had or could have had on regional and social economic development. This topic can be illustrated by the main conclusions of the ex-post evaluation of environmental measures co-financed by the ERD over the 2000–2006 programming period.

The main interest is to look at environment not as a constraint but as a driver for regional development. As part of an integrated strategy, environmental investments can have positive impacts on regional competitiveness factors (natural assets, economic environment, quality of public services, human resources) and offer opportunities to develop specific business sectors.

The evaluation of the use of the ERD for environmental infrastructures over the 2000–2006 period shed lights on this potential contribution. More than 20% of the ERD has been devoted to environmental measures, of which half to environmental infrastructures (mainly in the fields of waste water treatment and waste management). The main result was the improvement of public services in the fields mentioned above as well as some improvements in the environmental situation. Despite the relatively limited contribution to regional economic development, some interesting cases of integration of environmental investments into business development strategies have been reported.

The evaluation provides interesting findings which could be useful for the adaptation of the future role of the ERD towards promoting “green growth strategies” where the environmental sector could be a driving force.

Evaluating interaction of procedures and statistical indicators to address european regional programs towards results

*A. Dolce*

Regione Siciliana Nucleo di Valutazione e Verifica degli Investimenti Pubblici, Dipartimento Programmazione, Palermo, Italy

**Background**

“Convergence” is one of the priority Objectives to allocate the European Regional Development Fund (ERDF) resources under the 2007–13 programming period. Sicily, where the economic development still lags behind the average of the European Union (EU), falls under this Objective, which focuses on supporting sustainable integrated economic development and the creation of sustainable jobs. Sicilian Operational Programme (OP), aiming at modernising and diversifying regional economic structures, illustrates the strategic choices and priorities for ERDF spending, coherently with EU guidelines/regulations and based on the several evaluation processes conducted on the 2000–2006 previous programming period.

**Objectives**

In order to be strictly focussed on results at 2013, the OP is provided with extensive set of target indicators. The success of the strategy will depend on the distance from the target indicators at the end of the programming period. Hence the need for the Programme implementation procedures of taking into account, in some way and to some extent, these targets. This paper evaluates and measures the interaction between implementation modality of OP and its targets.

**Methods**

Reliability analysis, based on experts judgement to assess both the presence/absence of references to indicator and the distance reductions from targets, and association measures to establish the strength of the interaction, were applied to the 2000–2006 OP data.

**Results**

The results show that the interaction exists, in particular inside the evaluation criteria of operations: in presence of references able to properly select operations with clear impacts on indicators, significant reductions of distances from targets were observed.

**Conclusion/Application to practice**

To address a strategy towards results, selection criteria should include constraints to select operations with real impact on strategy indicators. These findings were used when redacting the OP which foresees that operations showing an impact on the correlated program indicator should be preferred.

**Keywords**: European regional development fund; Evaluation; Statistical indicators; Interaction; Implementation procedures
Evaluating the European Regional Development Fund contribution to climate change mitigation: the case of the Czech Republic

B. Lixon
ADE S.A., Louvain-La-Neuve, Belgium

During the 2000–2006 programming period, the European Regional Development Fund (ERDF) supported interventions in the field of climate change for a total budget of €2.3 billion. Nearly 70% of 2000–2006 Operationnal Programmes included at least a set of interventions aimed at mitigating climate change, i.e. environmentally-friendly, clean and economical energy technologies, renewable energy sources (RES) as well as energy efficiency, co-generation and energy control measures.

The Ex post Evaluation of Cohesion Policy Programmes 2000–2006 co-financed by the ERDF – Work package 5b: Environment and Climate Change showed that ERDF-supported interventions aimed at reducing GHG emissions and improving energy efficiency were quite exploratory. Some interesting results – the modernisation of heating systems, the increased use of RES and support to private investments for improving energy efficiency – were highlighted. However, evaluating the impact of such interventions in terms of GHG emission reduction still remains a challenge. The presentation will discuss a good practice in assessing such impact at programme level.

The Czech Industry and Enterprise Operational Programme implemented an innovative bottom-up step-by-step method to track achievements in energy savings and RES and to assess resulting GHG emission reduction. First, the selection of eligible projects was based on results of an ex ante energy audit including the assessment of avoided GHG emissions compared to a business-as-usual scenario. Second, the monitoring of co-financed projects provided information on installed capacity, energy saved and green electricity produced. An ex post assessment finally provided the estimated GHG emission reduction at programme level. Whilst data collection, control and processing are quite energy-intensive, this method remains reliable considering it is based on results of an ex ante energy audit confirmed by a compulsory monitoring.
Sharing lessons of experience

Best Practices in International Development Evaluation

T. Azzam¹, D. Stewart (Chair)¹, F. Carden², C. Lusthaus³, J. Raynor⁴, Z. Ofir⁵, N. MacPherson⁶, R. Conner⁷

¹Claremont Graduate University, Evaluation, Claremont, USA
²International Development Research Centre, Evaluation, Ottawa, Canada
³Universalia Management Group, Evaluation, Montreal, Canada
⁴TCC Group, Evaluation, Philadelphia, USA
⁵Evalnet, Evaluation, Johannesburg, South Africa
⁶The Rockefeller Foundation, Evaluation, New York, USA
⁷UC Irvine, Evaluation, Irvine, USA

This panel of international development evaluation practitioners will discuss best practices in evaluating Networks & Partnerships, Policy Influence, and Organizational Performance and aims to:

1. Identify a set of practices and principles that can guide evaluation in the philanthropy and development community.
2. Contribute to the conceptual knowledge and practice in the international development context.
3. Share the learning in these areas with other practitioners.

The following is topic outline:

Stewart I. Donaldson and Tarek Azzam’s presentation is titled New Challenges and Opportunities in International Development Evaluation. They will discuss some of the latest trends in this area and provide an overview and framework for the other presentations in this panel.

Jarod Raynor’s presentation is titled Networks & Partnerships Evaluation. She will argue that the evaluation of partnerships should focus on their intended purpose and ability to achieve their stated objectives. An evaluation could also examine their governance structure, financial management systems, and the transaction costs associated with their functioning.

Fred Carden’s presentation is titled Evaluating Policy Influence. He will argue that policy influence is often an incremental process requiring the correct combination of contextual factors and researcher actions to occur. To influence policy an evaluator needs to account for the stability of the government, foster trusting relationships with potential policy makers, and anticipate upcoming needs and interests.

Charles Lusthaus’s presentation is titled Organizational Performance Assessment (OPA). He will argue that advances in the OPA may be drawn upon in evaluations in order to address the growing needs of donor agencies, who wish to understand the performance and capacity of the organizations they fund.

Nancy MacPherson is the evaluation director at the Rockefeller Foundation and her presentation title is Donor Perspectives on Emerging Issues in International Development Evaluation.

Ross Connor and Zenda Ofir will serve as discussants for the panel.

Keywords: Developing Nations; Organizational Performance; Partnerships and Coalitions; Policy Influence; Foundations
Participants and co-production I

Participatory evaluation: who, what, why, how and when
C. Torrigiani, M. Palumbo, D. P. Congiu
University of Genoa, Di. S.A., Genova, Italy

This paper aims to propose some theoretical and methodological guidelines referred in particular to participatory evaluation processes. More specifically the Authors aim to shape a coherent and logical framework, from which some indications can be derived helping the evaluator:

– to select the social actors to be included in participatory evaluation processes, taking into account the costs and benefits arising from their participation (who);
– to suggest what issues it is more appropriate to discuss with the different types of actors included (what);
– to point out the specific objective facing participatory processes at different levels and at different stages (why);
– to identify the techniques and research tools more suitable for use in these processes (how).

All this, taking into account both the level (policy, program, project or service) and the phase of the policy cycle (strategy, design, implementation, conclusion) in which the evaluator is operating (when) both that participation in evaluative processes may be simply a means or a goal in itself, when it is aimed at the democratization of decision-making through the inclusion, dialogue and deliberation.

To focus more clearly and consistently these matters, it is particularly useful to adopt a theory-driven approach to evaluation: highlighting the cause-effect links between the components, functions and activities of the program, on one hand, and proximal and distal outcomes on the other, this approach helps the evaluator to understand, in the different stages of policy, which actors should be included in the evaluation, what issues should be discussed with them and which purposes should have these processes, especially in light of the fact that the uncertainty elements that characterize the decision-making and the implementation, elements that the evaluation should contribute to clarify providing useful information, take different weight and meaning at different times and for different stakeholders.

Keywords: Participatory evaluation; Theory-driven evaluation; Policy cycle; Decision-makers & stakeholders; Democratization

Co-production: a new concept in evaluation
P. Kettunen
University of Jyväskylä, Department of Social Sciences and Philosophy, Jyväskylä, Finland

The concept of co-production has only seldom been appearing in policy analytic and evaluation literature, and its implications on the evaluation theory remain largely open. As such, co-production seems to fit particularly well into the current emphasis on empowerment and user-oriented evaluation. Co-production as a concept does not necessarily limit itself to individual service-users; co-production by definition covers all kind of sharing in service production, either by organizations or by individuals. The emphasis in this paper will however be on individual citizens, who do their share as recipients of services or act in support of a particular policy through their own private activities (care of elderly by relatives, collection and recycling of waste, public safety for example). How do we conceptualize and evaluate such inputs?

One specific area where co-production is a necessary perspective is impact assessment. In many services the effect(s) of an intervention is highly dependent on the active involvement, the goodwill of the client. Co-production implies that even well-planned public interventions can fail if the requested recipient input is not adequate. The key to improvement is thus a better knowledge of the intervention and of the intended effects. This seems to be a well-known approach in the evaluation of teaching (PISA studies) but not generally practiced say in the evaluation of social work. For evaluation this means first that user-evaluation is emphasized more than before. Furthermore in impact evaluation the presence of the co-producer means that the net-result of an intervention is necessarily dependent on both the service provider and the recipient. Co-production can also be seen as a part of empowerment, i.e. that the wishes and preferences of clients get a bigger say.

Keywords: Co-production; User-evaluation
A participatory approach to developing evaluation capacity in the family resource program sector in Canada

P. Gabor
University of Calgary, Social Work, Lethbridge, Canada

Background
The purpose of this paper is to describe the process and products of a capacity building project, undertaken in association with the Canadian Association of Family Resource Programs (FRP Canada), a national organization representing over 2000 community-based programs.

Prior to the project, member programs had few resources and little expertise to conduct evaluations but were subjected to extensive demands for evaluation from funding authorities. The national association had thus identified building evaluation capacity as a priority.

Objectives
The main objective of the project was to build evaluation capacity, making it possible for local programs to respond to evaluation demands and for the sector to produce evidence about its scope and activities.

Approach
A participative approach was taken to assessing evaluation needs in the sector and to developing evaluation strategies and tools that could build capacity both at the local and at the sector levels. To ensure social validity, a nationally representative work group guided the process; frequent reference back to the field ensured the project’s products would be useful, accepted and utilized.

Products
The project resulted in a variety of tools and resources designed to help local organizations conduct evaluations and, as well, allow the national organization to aggregate data to the sector level. An on-line database was developed to complement the system, making possible not only on-line data collection but also real-time data analysis and reporting.

Outcomes
Currently, after two-years of implementation, over 400 program sites across Canada have used the system. Data from over 15,000 participants and 1200 staff members have been captured and the national organization has made extensive use of data to influence sector level programming and policies.

Keywords: Participatory evaluation; Capacity building; Sector evaluation; Empowerment evaluation; Community partnerships
Feeding evaluation results back into policy II

Theory based evaluation applied to policies

E. Canu, S. Pistis
Regione Autonoma della Sardegna, Evaluation Unit – CRP, Cagliari, Italy

Background
In the last 10–15 years, the EU has established in the Member States and Regions a common practice of evaluation, mainly based on indicators and case studies. Evaluation of Cohesion Policy has tended to be at the programme level, with a special focus on some measures. Recently, a tendency towards “strategic” evaluation has been witnessed. According to Council Regulation (EC) No 1083/2006 the purpose of strategic evaluation is “to examine the evolution of a programme or group of programmes in relation to Community and national priorities”). In Italy, the Strategic National Framework requires some institutional subjects (Regions and Ministries) to evaluate multi-funded policies as a whole. The scope of evaluative approaches has broadened to comprehend counterfactual and theory-based impact evaluation. Among these, Theory of change (ToC) is an approach aiming at understanding why an intervention produces intended and unintended effects, for whom and in which context. This approach is suited to the complexity of real implementation contexts.

Objective
In the US, ToC approach has been applied primarily at a programme level, but it could also be used at a policy one as well. To serve this purpose, the method should take into account all the different regulatory levels which affect public action both at the theory, and the implementation phase, since all the levels are connected to (possibly different and even contrasting) sets of objectives, priorities and instruments that gain the comprehensive end.

Result and application to practice
An enlargement of ToC method will be showed in a practical case based on an evaluative research of work services in Italy, a field in which European, national, regional and provincial decision making are implied.

Keywords: Theory of change; Policies evaluation; Methods; Evaluation of work services

Evaluation Synthesis for Policy Review

K. Fors¹; H. Schaumburg-Müller²
¹Andante – tools for thinking AB
²Copenhagen Business School

Background
In most policy fields evaluation has become institutionalised. In any given situation, policy makers have access to hundreds of evaluation reports – if they so want and if that information can be digested. Development cooperation is one of the most evaluated policy fields, with an estimated 3,000 evaluation reports being produced annually by the various multilateral and bilateral development agencies. The evaluation unit of one bilateral agency supplements its commissioning traditional evaluation reports with systematic studies synthesising global evaluation findings on development cooperation issues where the agency was rethinking its approach – and needed evidence to design new strategies.

Objectives
This paper discusses this experience of synthesis evaluation. It presents the aims and describes the methodological hurdles that face the synthesising process. The aim is to clarify the methodological options and present choices of one synthesis evaluation and to identify lessons learned.

Methods
The paper classifies the methodological challenges according to five overall issues; (1) definitions in a multicultural and multilingual environment, (2) sampling from an undefined and open population, (3) quality control in an area with no common standards and poorly developed quality criteria, (4) qualitative and quantitative methods of accumulating findings, and (5) reliability of the synthesis findings.

Conclusion/Application to practice
In retrospect the synthesis evaluation seems to be a good use of resources. It provides more strategic information in a shorter period of time and at lower cost than any single evaluation could provide. But practice of the synthesis evaluation builds on some assumptions around the availability and nature of evaluations and on the process of accumulating evidence. These assumptions can be right or wrong, and the necessary preconditions are not always at hand. The paper suggests when – and when not – synthesis evaluation are appropriate and how it can meet decision needs.

Keywords: Synthesis Evaluation; Policy Development; Methodology; Private Sector Development; Development Cooperation
Feeding evaluation results into policy-making: tools, timing, methodology and the case of the FP7 ICT interim evaluation

J. Gren, Dr
European Commission, DG INFOSO, 1000 Brussels, Belgium

Information and Communications Technologies (ICT) are woven into almost every aspect of our working and private lives and is essential to Europe’s competitiveness. To support competitiveness and to contribute to the Europe 2020 strategy for emerging from the current economic crisis, the ICT part of the FP 7 programme, with a budget of just over € 9 billion over seven years, is intended to be an important factor in building a single European Research Area (ERA) as well as in implementing the vision of Europe as a single, research-intensive and inclusive information space.

The ICT FP7 programme has recently been evaluated and the final report is due in May 2010. This paper will look at the use of this (interim) evaluation in feeding its results, recommendations and conclusions into policy-making for two purposes: 1) to provide guidance and steering (or reorientation) for the final part of FP7 and 2) to provide input to the design of any successor programme to FP7 as the preparations for FP8 will begin before the end of the current programme.

The paper will explore the use of evaluation results as a part of a social process, based on the interaction of individuals, decision-makers, stakeholders and the wider community but also between organisational methods, practices and routines. It will further explore the tools used to disseminate evaluation results, the methodology as well as the actual uptake of evaluation results and findings in subsequent policy papers or policy discussions. The concrete use of evaluation results, partly or fully, is clearly an important factor in effective policy-making.

Indeed, in the EU institutional set-up, evaluation is explicitly mandated to contribute to the design and implementation of policies mainly by generating relevant evidence in terms of attainment of objectives, production of results or impacts of various policies and programmes.

**Keywords:** Evaluation results; Policy-making; ICT FP7; Use of evaluation results

Evaluating ICT Research in the 7th EU Framework Programme

M. Übelhör, C. Oliveira
European Commission, DG Information Society and Media, B-1049 Brussels, Belgium

**Background**

With about 1.2 bn EUR of investment every year, ICT research is the biggest part of collaborative research funding in the EU Framework Programme.

**Objectives**

It was the purpose of this interim evaluation (programme duration 2007–2013, evaluation conducted 2009–2010) to assess the quality of research, the quality of the implementation of the programme and the overall progress towards the programme objectives.

**Methods**

European Commission conducted this evaluation with the help of a panel of independent experts as well as a team of professional consultants. This presentation will present the evaluation process from the perspective the Commission. The methods used were, inter alia, surveys, expert interviews, portfolio analysis and composition analysis. The evaluation drew on a number of further studies, including ones using bibliometrics and network analysis.

**Results**

The results showed that the quality of the research is good and that the programme is well-implemented in its topical orientation. However, the expert panel found that the administrative costs of proposing and participating to the programme were too high.

**Application**

The insights and judgments from this evaluation will feed into future policy-making in various ways. These mechanisms are foreseen to be subject of another presentation (proposal submitted by Jörgen Gren, European Commission).

**Keywords:** RTD policy; European Union; Framework Programme; Innovation policy; Evaluation
Evaluating the complex

Presentation: methodological development in complex evaluation: focussing evaluation questionsCase of the country level evaluation Ethiopia 2004–2010

J. Chomel
European Commission, Brussels, Belgium

The formulation of evaluation questions is the first major element to get a good quality evaluation report. The Joint Evaluation Unit of the External relations family in the Commission of the European Union has been doing so now for almost ten years and accumulated experience in about 50 Country Level Evaluations. Over time and learning lessons, the art of asking evaluation questions has evolved.

The presentation compares questions of the two evaluations, one finalized in 2004 and one ongoing in 2010 concerning EC cooperation with Ethiopia. It demonstrates how focussing questions allows for deeper analysis at a specific point in the results chain. Linking questions to the intervention logic and putting them at the level of outcomes or impacts reinforce validated insights for impact-oriented evaluations.

Lastly the presentation proposes best practices in formulating evaluation questions

Keywords: Evaluation methodology; Ethiopia; Evaluation questions

Making sense of Sensemaker®: Evaluating development initiatives through micro-narrative capture and self-tagging in Kenya and Latin America

I. Guitj1, J. Hecklinger2, M. Maxson2
1 Learning by Design, Randwijk, Netherlands
2 GlobalGiving, Washington DC, USA

In evaluation, ‘stories’ and ‘statistics’ are often posited as an either/or choice, and when combined, their inherent limitations persist. Statistics ignore outliers, are without context, lend themselves with difficulty to explanation, and focus on the knowable. Stories convey few detailed voices, lend themselves with difficulty to representative insights, and researcher bias introduces bugs to both.

What if the best of both worlds were possible? Sensemaker® offers that potential. Through ‘mass capture’ of micro-narratives, self-indexing, and rapid, sophisticated statistical analysis, patterns emerge and become the basis of analysis, thus reducing researcher bias and generating real-time insights.

For both GlobalGiving and RIMISP, measuring extent and complex impact of their programs offers evaluation challenges. GlobalGiving has created an online marketplace where organizations worldwide seek funding for thousands of specific projects and thousands of donors worldwide provide that funding. RIMISP works in 11 countries in Latin America with over 170 organizations on applied research and policy to generate rural alternatives within territories, nationally and internationally. Traditional evaluation methods would be prohibitively expensive and would not necessarily capture emerging impacts.

Firsthand experience of using Sensemaker® is discussed for these cases, involving stories gathered from community members, researchers, and social change agents, enabling the visual depiction of patterns of change. Comparing how the signification framework emerged, story collection and self-indexing processes, and analytical value adding will give insight into what Sensemaker® offers the evaluation world.

Rooted in complexity theory, which looks at systems that are inherently unpredictable and cognitive science, which considers how people make sense, this experiment explores how multiple perspectives illuminate underlying patterns when more traditional means of evaluation are not workable. These experiences illuminate a path for others seeking scalable and cost-effective ways to make sense of many and diverse initiatives covering wide geographic areas.

Keywords: Narratives; Complexity; Rigour; Self-signifying; Kenya and Latin America
Towards a complexity oriented monitoring and evaluation practice

J. Van Ongevalle1, A. Maarse2, H. Huyse3, C. Temmink4

1 HIVA, Development Cooperation, Den Haag, Belgium
2 PSO, Capaciteitsopbouw, Den Haag, Netherlands
3 HIVA, Development Cooperation, Leuven, Belgium
4 PSO, Capaciteitsopbouw, Den Haag, Netherlands

The recent financial crisis and the growing demand for better results in the development cooperation sector are just a few examples for the need to rethink the underlying paradigms that inform our practice of development cooperation. Until now these have been determined by a linear reductionist way of thinking characterized by a strong belief in predictability of social processes. At the same time, donors and policy makers still face problems to receive all the necessary information about concrete results and effectiveness from the programmes they fund. This is especially the case in complex programmes of social change that are faced with unpredictable outcomes and where a multitude of actors and factors contribute to the results.

Complexity theory is a fast growing field that offers great potential to address these challenges. In this paper we will narrate how complexity oriented monitoring and evaluation methodologies such as outcome mapping and most significant change have been piloted by 15 international NGOs in different country programmes through a longitudinal participatory action research process. The paper provides a thorough description of the research methodology and reports on the results of the first year of the study. Through a cross case analysis we explore the effectiveness of the monitoring and evaluation systems in promoting collaborative learning between programme implementers and beneficiaries, evidence based planning, improved accountability and sustainable programme outcomes. Furthermore, a framework is developed that will help programme staff and their stakeholders to choose the appropriate monitoring and evaluation system for a specific context. This is done through a detailed contextual analysis of each case in the action research.

Keywords: Complexity; Action-research; Outcome mapping; Most significant change
Evaluation in education

School effectiveness and policy context. Using evaluation and comparison for understanding and improving the Italian education system

F. Cornali
University of Turin, Social Science, IT10133 Torino, Italy

The evaluation of how well institutions of education perform has been the topic of increased attention in recent years. The key role of competence and knowledge in stimulating civic development and economic growth has been widely recognised by several studies and researches. As a consequence, a number of indicators purporting to measure various aspects of the effectiveness of education system are published frequently.

International collections of comparative statistics and reports in the field of education — e.g., the editions of Education at a Glance by the Organisation for Economic Co-operation and Development (OECD) — show that in Italy considerable advancements has been made in recent decades. Despite this progress, however, the Italian education system efficiency remains low in producing quantitative outputs — e.g., the Italian enrolment rates in upper secondary and tertiary education is lower than the average across OECD countries — and in producing qualitative outputs — e.g., in mathematics the average knowledge level of students aged 15 is low in international comparisons, since Italy ranks among last positions in the Programme for International Student Assessment (PISA). Not only, the Italian education system reveals inefficiency in using resources.

This paper, using heterogenic sources (national and international statistics, norms and laws), describes the main features of the Italian institutions of education (with special attention to two case studies representing territorial differences in Italy: the province of Trento and Sicily) and analyses the education policies.

The aims is underline the relevance of comparison and evaluation to define effective education policies. Through international and interregional comparisons, countries may come to recognise strengths and weaknesses in their own systems, but only through a constant monitoring and evaluation it is possible to assess important features of the functioning and impact of education under the specific conditions and environments in which education systems operate.

Keywords: Evaluation; School improvement; Italy; Comparison; Education policies

Evaluation of the implementation of an education program in order to control obesity in Rio de Janeiro.

E. M. Dos Santos, M. M. Cruz, A. Leal
Fundacao Oswaldo Cruz/ENSP, densp, Rio de Janeiro, Brazil

The changing nutrition’s landscape of the Brazilian population shows that while declines the incidence of sub nutrition in children and adults are at accelerated pace, the obesity rates increases. This study aims to evaluate the implementation of an education program in obesity prevention and control of in Rio de Janeiro, Brazil. The study of multiple cases. Two units were selected considering the implementation time at each health care facility, the number and the contrasting social-economic condition between two groups of clients assisted by the health facilities. It was a mixed methods evaluation and judgment was performed considering technical components and program theory. Technical components of the program and structure of the program were assessed for compliance using the state guidelines as technical standards. The health facility A (HFA) and health facility B (HFB) had their degree of implementation classified according quartile intervals. Program theory and the pedagogical strategy were assessed contrasting direct observation and health professionals focus in depth interview with the international and national literature on health behavior change theory and transformative education theory. Judgment was done taking in intervention in its context and the specified evaluation dimensions. Results show that in both health facilities, the intervention is partially implemented, and in both services, issues related to infrastructure are the major problem for the implementation. Barriers to implementation in HFB also include issues related to work organization and planning. Regarding program theory and the pedagogical strategy we conclude that in both cases participatory and interactive educational approaches were compromised by physical space and continuity of the intervention. In HFA and HFB the implementation of participatory educational approaches, based on dialogue and interaction between clients and the health professional team is constrained by the context and the structures of services organization.

Keywords: Obesity; Nutrition; Education; Evaluation
Climate change and Educational change – Evaluation and Public policy in Greenland

M. W. Bolesen
University of Copenhagen, Department of Political Science, Copenhagen K, Denmark

Background
In 2005 the Greenlandic Home Rule government put forward a proposal to strengthen the educational sector – qualitatively and quantitatively. In 2006 it was arranged that the Educational Reform was to be monitored and evaluated with the intent to identify special areas and/or special means in order for the development to take place in the best way. A program evaluation is taking place.

Objectives
The overall objective is to increase education in the Greenlandic population. In 2005 only 1/3 of the adult population had a formal education above elementary school. The objective is to increase that proportion by 2020 to 2/3.

Methods
The government is spending lot of money in the educational sector. And they have received/are receiving approximately 28 million € per year from the European Commission. (Greenland stopped being a member of the European Union in 1989).

Results
In the Educational Reform the work is focused on several areas. Among the most important ones you can find: organised development and educational development in order to support civil society; development in isolated parts of the (huge) country; qualified participation in globalised society; increasing geographic and social mobility; strengthening of cultural identity; change in educational arenas; change in the role of the teacher; changes in possibilities for (more and higher) education. During the first few years the results were ‘promising’ – more and more students in the educational sector. But during the past year it looks as if the educational market has become too popular; the other sectors of society can not keep up.

Conclusion/Application to practice
This poses serious and unpleasant questions: what has happened? why did it happen? what can be done about it?
Evaluation of water related interventions

Roundtable on the evaluation of water-related activities in developing countries

M. Sieburger
KfW Entwicklungsbank, FZ E, Frankfurt am Main, Germany

The latest report of the Independent Evaluation Group of the World Bank, “Water and Development: An Evaluation of World Bank Support, 1997–2007”, examines the World Bank’s water-related activities in developing countries and makes recommendations for an improved support of countries in managing their water resources against the backdrop of growing water shortages, pollution and climate change. Inspired by this extensive evaluation KfW development bank’s evaluation department and IEG are currently organizing a conference aiming at bringing together the results of thematic evaluations with a similar focus as the IEG study. The conference will be held on August 31st and September 1st 2010 in Berlin and presumably will center around the topics of water, decentralization and waste water treatment in developing countries. So far participants and topics are not finally determined. The intention is to bring together a limited number of practitioners from evaluation units who have recently carried out evaluations of activities within the thematic focus mentioned above, to present proceedings and results of their studies and to exchange experiences on what methods worked, about lessons learned and to discuss which evaluation results brought about change and why.

Our suggestion is to present and discuss the findings of the joint conference of KfW and IEG at the European Evaluation Society International Conference and therewith address a greater audience and foster exchange. Within a roundtable of participants evaluation methods and possible improvements can be further discussed and the value of evaluation of water-related fields to developing countries under growing water stress considered. At least two of the participants of the Berlin conference would present brief introductory reports.

Keywords: Water supply; Waste water; Sanitation; Decentralization
Evaluation of systems and findings at the discretion of the policy makers

Evaluation research and public administration: a regional observatory on evaluation and quality of services

A. Calabria, A. Rissotto
CNR - National Council of Research, Institute of Cognitive Sciences and Technologies, Rome, Italy

The objective of this contribution is to study the relationships between evaluation and politics analyzing the "Observatory on Evaluation and Quality of Services".

The Observatory was implemented by the Institute of Cognitive Sciences and Technologies (National Research Council) and funded by the Regional Customer Protection Authority of Lazio with the aim of involving citizens-users in the evaluation of services.

The project was developed starting from the idea that quality of services is a right and that evaluation is a negotiation process among legitimately different points of view. Meanings attributed to the quality are linked to the expectations, needs, interests and mandate of the different stakeholders and are related to a specific temporal, social and cultural background.

The Observatory is an interface between citizens and institutions and has been conceived as a space for interaction among the Regional Authorities, the Consumer Associations, the Service Providing Organizations and the Citizens-Users.

The Observatory's research activity involved the evaluation of several services and intervention systems (public transport, pharmaceutical assistance, reservation of health care services, emergency health services) and the study of phenomena (commuting and poverty) having an effect on the citizens' quality of life.

An analysis of the principle strengths and weaknesses that emerge during the implementation of the project provide a key to explore the relationships between evaluation and politics. The principal strength was the choice of research themes of significance for the population and cutting across the different policies (sustainable mobility, housing, social and economic policies). A further positive element was the direct involvement of citizens in the research.

The weaknesses included general aspects such as the critical nature of the relationship between evaluative and administrative culture and specific aspects such as the relative lack of interaction among various stakeholders.

Keywords: Quality of life; Consumer protection; Citizens' rights; Public administration

How is evaluation accommodated in emergency policy-making:
A comparison of post 9/11 emergency legislation in the UK and the US

M. Fazekas, C. van Stolk
RAND Europe, Evaluation and Audit, Milton Road, United Kingdom

There is a common held assumption that principles of evidence-based policy making struggle to meet policy-making demands in turbulent times. In terms of evaluation, it seems logical to assume that as decision makers need to find answers to emergent policy issues in difficult times the normal stages of policy cycle are less relevant and applicable. Therefore normal evaluation mechanisms are likely to be used less.

This paper is not concerned with how decision makers arrive at decisions or use evaluation to shape decisions in difficult times. This paper aims to look at how evaluation is or is not accommodated within emergency policy making. Though it may be true that evaluation in these times matters less to decision makers to arrive at policy answers, the relative accommodation of evaluation in the policy itself can give an indication of the importance of evaluation and accountability in the wider policy making process and also the willingness of policy makers to return to the normal cycle as soon as possible.

The paper aims to look the introduction of emergency policy after 9/11 in the US, and the UK and assess how evaluation was or was not accommodated within the policies introduced. The chapter will first outline what the evaluation arrangements were in emergency policy and where these arrangements were accommodated within existing evaluation frameworks. Subsequently, the paper will outline what explains the differences between countries in terms of a number of factors, such as existing evaluation culture and arrangements, the perception among policy makers and public of the severity of the crisis, the need for accountability from parliamentary and other overseers, etc. Finally, the chapter will discuss what this comparison can tell us more broadly about how evaluation fits within emergency situations and also how policy makers use evaluation for different reasons in emergency policy.

Keywords: Evaluation in emergency policy-making; The politics of evaluation; Policy cycles

“Drafting an International Evaluation Policy – Lessons from IFRC”

S. Chaplowe
International Federation of Red Cross and Crescent Societies (IFRC)

As a leading global actor in the humanitarian field, the International Federation of Red Cross and Red Crescent Societies (IFRC) recognizes the importance of ethics and transparency in its work. Towards this end, IFRC has recently drafted an evaluation policy that combines standards that guide how evaluations are conducted with criteria of what is evaluated. This presentation will examine the key challenges and considerations in drafting a policy for an organization with a membership of 186 national societies. It will summarize the content of the policy, highlighting how it attempts to present more than just standards and criteria, but specific procedures to ground them in practical practice. The IFRC policy is an interesting case study of a large, international organization adopting a policy to promote reliable, useful, ethical evaluations that contribute to organizational learning, accountability, and its mission to best serve those in need.
Methodological challenges II

Methodologies for monitoring and evaluation in multi-actor settings – A case of regional water management in the Netherlands

L. Hermans
Delft University of Technology, Faculty of Technology Policy and Management, Delft, Netherlands

Policies are increasingly formulated and implemented through collaborative processes that involve various actors. These participatory policy processes bring new challenges for evaluations. Simply focusing on one specific actor and its needs is not likely to bring about the collaboration of others, which is needed for meaningful evaluations of these collaborative processes.

This presentation will discuss a case in the Netherlands where a water management plan has been developed in a collaboration between a regional water board and its partners, as part of the implementation of the EU Water Framework Directive. A framework to monitor and evaluate the implementation of this plan has been developed, combining three methodological building blocks. First, a theory-based evaluation approach has been used. Second, insights from actor analysis have been used to ensure that the views of different actors are taken into account, building cognitive maps that represent causal chains as perceived by different partners involved. Third, elements of adaptive policy analysis have been used to identify critical assumptions in these cognitive maps, along with so-called signposts and triggers, which help to identify the main indicators that need to be monitored. The resulting framework should help not only to monitor plan implementation and goal realization, thus contributing to official evaluations required under the Water Framework Directive, but should also enable plan adaptation during implementation in the period until 2015.

The combined use of these methodological elements proved feasible and provided useful insights into the content of the plans, and its monitoring. Also, the focus on policy monitoring, forced policy analysts and evaluators to think about plan implementation, an issue that was not normally at the forefront of their attention. Having an instrument that helps focus the attention of analysts and planners on this aspect of a policy process, seems worthwhile in its own right.

Keywords: Methodologies; Multi-actor systems; Water; Monitoring

Methodological challenges in the evaluation of social and public health programs from a qualitative paradigm

A. Farre ¹, M. Pia ¹, M. Garcia ², R. Crespo ¹, A. Cardona ¹, N. Codern ¹
¹ Autonomous University of Barcelona, Chair in Qualitative Research, Barcelona, Spain
² Catalan Institute of Oncology, Cancer Prevention and Control Unit, L’Hospitalet, Spain

Background
The starting point for this presentation is the evaluation of the Colorectal Cancer Screening Program, implemented by the Catalan Institute of Oncology. The evaluation focused on the application of the program in two municipalities, Vilafranca del Penedès and Hospitalet de Llobregat, both from the area of Barcelona, Catalonia (Spain).

Objectives
This presentation raises methodological issues concerning evaluation from a qualitative standpoint and the utilization of data analysis techniques like Discourse Analysis (DA) or Thematic Content Analysis (TCA). As well as it sets up a starting point for considering the contribution of the Frame Analysis technique to the evaluation of social and public health policies, programs and projects.

Methods & Results
The main goal of the data collection phase of the evaluation was to explore the perceptions of the people invited to take part in the Program, and specifically about their own reasons for participation/non-participation or adherence/non-adherence to it, as well as their views on what can be improved. In order to get this data, 11 focus groups and 4 triangular groups were carried out, involving a total of 78 people from both municipalities. At the end of the process, several recommendations for change were done, regarding the design and the processes underlying the Program, in order to increase participation and adherence to the Program.

Conclusion/Application to practice
On the basis of this evaluation experience, several methodological challenges of evaluating social and public health policies, programs and projects can be illustrated, i.e. the crystallization of the dialogue between the formal theories underlying the design of policies/programs and the informal theory impregnating the practices and discourses (from users and professionals) which is emerging from the qualitative field work.

Keywords: Qualitative methodology; Social programs; Public Health; Users perspective; Frame Analysis
Qualitative evaluation of a non-contributory retirement system in México

M. d. P. Torres Pereda¹, A. J. Flores², M. Cervantez³, M. M. Téllez-Rojo⁴
¹ Instituto Nacional de Salud Pública, CISS/CIEE, Cuernavaca, México
² Universidad Autónoma del Estado de Puebla, Antropología, Puebla, México
³ Instituto Nacional de Salud Pública, CIEE, Cuernavaca, México

In México there is an elderly population of 7.9 million. In rural areas, 50% of them are living in poverty. However, existing in-country cash transfer (CT) programs that have been evaluated showing positive results leave aside the poor elderly and do not evaluate the perceived impact from the beneficiaries’ perspective. The Mexican government has created a National Elderly Support Program; a non-contributory retirement system that aims to help elderly adults living in poverty not only in terms of income but by strengthening social safety nets too. By 2009, two million elderly were beneficiaries, receiving 50 USD monthly.

Objectives: To evaluate the effects achieved by the program on the living conditions of beneficiaries. In particular, we aimed to explore the perceived impact of the program by beneficiaries.

Methods: This is a quantitative – qualitative evaluation. The qualitative team visited four communities (two were indigenous; two were non-indigenous). Semi-structured interviews (N = 136) were carried out with the elderly, and four non-participatory observations of the payment day were evaluated. We present qualitative results for perceived impact on quality of life, specifically findings on emotional health and social networks.

Results: Improved emotional health was consistently reported and conclusive using both methodologies. This impact was strongly expressed by women and related to the security of having a regular income, reducing food insecurity, and reducing economic dependence. A greater activation of their social networks due to greater reciprocity of goods and food was reported especially among the indigenous. CT was not used exclusively in the elderly but by all members in the household.

Conclusions: The perception of users should be considered as part of routine impact evaluations. Targeted programs for poverty alleviation should be assessed by understanding cultural and structural contexts as the impact can be seen in other household members and not only on the direct beneficiaries.

Keywords: Qualitative evaluation; Perceived impact
Joint Evaluation

Joint Evaluations – experiences from international financial institutions – the cases of EBRD, EIB and IFC.

W. Gruber1, S. Edelmann2, W. Schmidt3
1 European Bank for Reconstruction and Development, Evaluation, London, United Kingdom
2 World Bank Group, IEG, Washington DC, USA
3 European Investment Bank, Operations Evaluation, Luxembourg, Luxembourg

Independent evaluation units of International Financial Institutions (IFIs) represented in the Evaluation Cooperation Group (ECG) increasingly strive to engage in joint evaluation work, where meaningful and feasible. But, despite the emphasis of MDBs and expectations raised in the 2005 Paris Declaration on Aid Effectiveness on completing joint evaluations, their number has been relatively limited.

From a practical perspective there are numerous potential benefits and real challenges, such as:

• Potential benefits: harmonisation of approaches, procedures and ratings; reduction of transaction cost and ‘client fatigue’; potential for cross-leverage effects; mutual learning and capacity building; widening of evaluation scope; potential for increased objectivity, transparency and independence;

• Real challenges: varying mandates, institutional histories, cultures, systems and procedures not allowing inter-IFI standardisation of evaluation frameworks; Differences in project conditionalities and covenants; incongruencies in project priorities and maturities; requirement for inter-IFI coordination and longer lead-time and provisioning for consultative processes.

This session will (a) provide further background on why international institutions encourage joint evaluations; (b) present benefits and challenges of joint evaluation work, (c) provide examples of different types of approaches to completing joint evaluations; and (d) discuss real cases where EBRD, EIB and IFC have conducted joint evaluations.

Keywords: Joint evaluation; IFI

Joint Donor Evaluations of Conflict Prevention in Democratic Republic of Congo and in Sudan

E. Brusset
Channel Research

Objectives
This paper will capture the key learning from the implementation of two large multi-donor evaluations of conflict prevention in peace-building in two developing countries (South Sudan and DRC). These evaluations included a steering committee of 11 to 18 members, and participation by the national authorities. They touched on sensitive national issues but attempted some innovative methods in focusing on national politics and impact.

Methods
The methods used were the establishment of reference groups, extensive conflict analysis, portfolio reviews including access to sensitive national documentation, and project based impact assessment using a sample of case studies.

Results
Principles for improvements to the conduct of Joint Donor Evaluations in developing countries will be proposed, as well as for the inclusion of national authorities, and on the management of sensitive information. The presentation and paper will also provide considerations on impact assessment in unstable environments, and the specificities of evaluating conflict programming.

Conclusion/Application to practice
The OECD DAC is currently piloting some draft Working Guidance on the Evaluation of Conflict Prevention and Peacebuilding. This presentation will address some of the aspects of this guidance, with a view to finalizing it. The presentation will also provide some insights into large joint evaluations. Finally it will outline the key challenges for the future in the evaluation of stabilization and peace interventions in complex crises.

Today many organisations are struggling with the complexity and unrealistic nature of social reporting, with the blurred notions of responsibility and attribution that these entail. This presentation will show that evaluation is a full fledged and legitimate form of stakeholder engagement that can complement planning and coordination.

Keywords: Methods; Conflict Prevention; Joint Multi Donor Evaluation; Congo and Sudan; Lessons Learned
The political economy of joint evaluations

**O. Winckler Andersen**, E. Broegaard

1 Ministry of Foreign Affairs of Denmark, Head of Danidas Evaluation Department, Copenhagen, Denmark

2 Ministry of Foreign Affairs of Denmark, Evaluation Department, Copenhagen, Denmark

**Background**

Both in the Paris-Declaration and the Accra Agenda for Action donor harmonization is pursued. Also within the evaluation community joint evaluations receive increasing attention, and are generally expected to result in lower transaction costs on both donor and recipient side and to enhance relevance and use of evaluation-based knowledge. In recent years various assessments of harmonized evaluations have been published, but the actual advantages of joint evaluations are rarely questioned.

**Objectives**

To critically assess the assumptions and implications of harmonized evaluation efforts by applying an analytical framework that focuses on transaction costs and incentives; and point to factors that may foster a more focused and effective approach.

**Methods and Results**

The paper will outline a typology of joint evaluations to help distinguish some of the features that influence the possible costs and benefits of joint evaluations. Drawing on this typology and relevant examples, the paper applies an analytical framework based on new institutional economics and principal-agent theory with focus on transaction costs, incentive structures, etc., to explore the political economy of harmonized evaluations. Amongst issues explored are:

- Economies of scale versus increased complexity and analytical gain;
- Coordination issues and informational problems;
- Significance of the interests of actors involved.

The paper finds that due to a number of common characteristics for harmonized evaluations, e.g. evaluation object and incentive structure, there is a risk that complexity increases to the point that it poses a risk to evaluation process and product, and that only in specific circumstances harmonized evaluations will lead to the anticipated advantages.

**Conclusion/Application to practice**

The paper concludes that fuller consideration of the possible costs, benefits and interests in play when conducting harmonized evaluations are needed, and points to how a more targeted and strategic approach may enhance the chances of reaping the full rewards hereof.

**Keywords:** Joint evaluation; Transaction costs; Complexity; Incentive structures; Political economy analysis
How independent are evaluators and evaluations?

How independent are evaluators and evaluations?

W. Meyer¹, R. Stockmann¹, N. Stam², E. Monnier³, J. E. Furubo⁴, H. Cummings⁵

¹Saarland University, CEval, Saarbrücken, Germany
²University of Rome, Italy
³Euréval Center for European Expertise and Evaluation, France
⁴Audit-Director at the Swedish National Audit Office, Sweden
⁵University of Guelph, Environmental Design and Rural Development, Guelph, Canada

The main task of evaluations is to produce scientifically sound results of an independent research. Evaluation standards are willing to define more precisely what independence of evaluations is about. The DAC-Principles, for example, highlight the independence of evaluations from the policy-making process and the operational departments within an implementation agency, while the JCSEE-Standards focus more on transparency and a fair and open communication process. In public discussions, the trustworthy of evaluations is sometimes doubted because of the dependencies produced by the contract between evaluators and their clients, assuming evaluation results per se biased by the interests of the contractor. While talking to Evaluators, some report about trials to influence their results or interpretation more or less deeply, sometimes accompanied by pressure or even threats. On the other hand, principals mention the market dependency and the interests of evaluator not to spoil the relationship by uncomfortable conclusions, leading sometimes to poor reports with little evidence for practical reforms.

There is still a lack of information on these several aspects of independence and its importance for evaluation practice. As an input to the planned panel debate, results from a recently conducted survey among the members of DeGEval on this topic will be presented. The results will be commented by members of other European Evaluation societies by using experiences from their countries to highlight similarities and differences.

**Keywords:** Independence of Evaluations; Evaluation Ethics
Evaluation, cohesion and integration

Territorial impact assessment: A theoretical policy instrument or political theory building?

M. Toudal
Centre for Regional and Tourism Research, CRT, Nexo, Denmark

In the recent years a new discourse focusing on the territorial cohesion of the EU has come into fashion in the policy field of regional development and cohesion policies, most notably with the “territorial agenda of the EU” and ESPON Programme (European Observation Network for Territorial Development and Cohesion) which is supposed to support policy development related to EU Cohesion Policy. In the same time efforts to strengthen the use of evaluative policy instruments in the EU have been introduced as different “impact assessment” tools. These developments have met in the ambition to construct a tool or instrument for “Territorial Impact Assessment” (TIA).

On basis of the research based project “EUROISLANDS” under ESPON, this paper analyse the theoretical and political challenges to develop a framework for territorial impact assessment specifically for islands. The theoretical challenges are restaging the traditional methodological and conceptual problems from the evaluation literature in a new context. The classic questions of use of the evaluation results (and methods) are also present in the close cooperation with stakeholders with interest in the regional policies concerning islands.

The paper analyses the challenges outlined above and sketch the results from the work in the EUROISLANDS project. No one with the faintest knowledge of evaluation will be surprised that politics are important in evaluation; the research question of this paper is therefore how different interests and institutions have influenced this area of regional policy in the EU. The more general questions the results can help to understand is hopefully how and why evaluation play such a significant role in the regional and cohesion policies of EU.

Keywords: Impact assessment; Territorial cohesion; Regional Policy; Stakeholder involvement

From Candidacy to Cohesion – Lessons from recent thematic evaluations by the European Investment Bank

G. Dunnett
European Investment Bank, Operations Evaluation, Luxembourg, Luxembourg

Support for EU cohesion policy is one of the main operating priorities of the European Investment Bank, and the EIB regularly evaluates its lending activity in this area based on the detailed evaluation of a sample of individual operations. In addition to looking at the support given to member states (most recently in Germany, Ireland and Spain), recent evaluations have also examined the role of the EIB in the pre-accession process, both for new member states prior to their accession and for candidate and potential candidate countries. The paper will also present some initial findings from the ongoing evaluation of the EIB contribution to the JASPERS initiative aimed at providing TA support to new member states preparing major projects during the current programming period.

The paper will outline the scope and methodological approach used for these recent evaluations and will compare the findings, drawing out common issues relating to the difficulties facing countries making the transition from pre to post accession support and their potential implications for cohesion policy.

Keywords: Cohesion; Pre-accession

How the use of evaluation laid the groundwork to facilitate European integration

G. Ellis, M. Pai
Gells Communications, Strategic communications, 1050 Brussels, Belgium

The introduction of euro banknotes and coins in 16 European Union Member States since 2002 constitutes a major landmark for European economic and political integration. The intent is to bring as many of the European Union’s 27 members into European Monetary Union as possible.

One of the European Commission’s Directorate-General for Economic and Financial Affairs (DG ECFIN) responsibilities is to communicate on the euro which it does largely via its website, and by making brochures and expert speakers available in countries which aim to adopt it as their currency. However, DG ECFIN was concerned that the subject of European Monetary Union, the euro and the euro the European Commission plays in this area, was not recognised in the schools curricula across Europe. Therefore, DG ECFIN hired Gells Communications to perform an evaluation of the status quo, and to provide conclusions and recommendations on how the subject of the euro and the European Commission could be introduced to support existing curricula across Europe.

The study involved extensive desk research, engaging pedagogical experts, six country-level focus groups involving 112 teachers, and two online consultations in over 20 countries with responses from approximately 1,300 teachers. Over the course of this study, Gells produced an evidence-based strategy which outlined the process needed to develop and disseminate pedagogical resources across 27 EU Member States, providing indicators to evaluate the level of success of such an initiative.

The next phase for DG ECFIN was the development of pedagogical resources, and a dissemination strategy for these tools and resources. The aim of this is to facilitate the introduction of the euro and EMU to support existing schools curricula across the EU 27 Member States. DG ECFIN fully accepted Gells Communication’s recommendations and is currently developing pedagogical materials.

Keywords: DG ECFIN; Strategic communications; Education; Multi dimensional communications; Evaluation
Impact evaluation I

Learning from impact evaluations – Findings in the area of the European Social Fund

J. Toulemonde1, D. Reyntjens2
1 Eureval and Lyon University, Partner, Lyon, France
2 European Commission, DG Employment and Social Affairs, Brussels, Belgium

Impact evaluation is gaining momentum and raises a lot of expectations. In this context, the European Commission has investigated into the practice and use of impact evaluation in the case of the programmes supported by the European Social Fund (ESF). The study was undertaken in co-operation with several Member States (Belgium, Hungary, Italy, and Poland) having carried out impact evaluations over the last years.

All reviewed evaluations delivered quantitative estimates of the impact of training programmes targeted at unemployed people. Quasi-experimental methods were applied with various types of control groups, diverse analytical approaches, a range of different findings, and uneven utilisation.

This communication follows the logic of the study which tested a series of assumptions about the reviewed impact evaluations, i.e.

- They addressed issues that were relevant to decision-makers
- Their findings were timely, valid, and trusted
- They fed into some decision-making processes

Building upon this test, the authors reflect on how to learn more and faster from impact evaluations in the context of the new ESF programming cycle due to start in 2014.

Keywords: Utilisation; Impact evaluation; Transferable lessons

Why so few impact evaluations?

P. Severati
Isifo – Educational Development Institute, Esf Evaluation Unit, Roma, Italy

Impact evaluation is more and more considered in official documents and programs. Nowadays many researchers are able to carry out counterfactual analysis and several statistical packages are affordable or freely downloadable from the web. Large user’s communities help answering technical questions about routines, procedures, controls to solve analytical problems. Yet, especially in countries where impact evaluations started to develop under the impulse of the Structural Funds Regulations, rigorous evaluations regarding the effects of public policies are very rare.

The question the paper tries to address is the following: given the above premises, why so few impact evaluations? What are the major difficulties in order to carry out good impact evaluations? As well known, evaluating the effects of public policies requires the adoption of sophisticated methods of analysis in order to treat selection and self-selection problems. Unfortunately, only few people are really aware that building up complex information systems and data sets is a major obstacle.

This paper sustains the following thesis. Monitoring first and then evaluation require a strong institutional consensus regarding their scope and objectives, not only on the part of policy-makers but also on the part of other important and often neglected actors. Broad monitoring and evaluation characteristics should be known at administrative and operational level in order to gather good information on beneficiaries or on the way interventions are implemented by agencies. Data are not freely disposable in nature but are the product of a collective learning process which requires shared knowledge. Good data will be available only when everyone involved in this process knows that it will be used for research in order to understand if public policies work or not.

The paper describes how these problems were tackled in Italy in the context of the European Social Fund (ESF).

Keywords: Impact evaluation; Counterfactual analysis; Data collecting; Evaluation and information demand; Shared knowledge
Real time evaluation

R. Polastro
Fundacion DARA Internacional, Madrid, Spain

In recent years, efforts have been increasingly directed toward improving humanitarian response through real-time evaluations (RTE). An RTE can be defined as an evaluation carried out at the early implementation states of a humanitarian operation which almost simultaneously feeds back its findings for immediate use by the broader humanitarian community at the field level. An RTE is primarily intended for sudden-onset disasters, or protracted crises undergoing a phase of rapid deterioration or escalating violence. These evaluations differ from other forms of humanitarian evaluation regarding speed, coverage, methods, and outputs. RTEs are typified by their shared management and methodological oversight through global and national level inter-agency reference and management groups; celerity of mobilization, feedback and follow-up; light, agile approaches; restricted scope; and participatory methods. Ideally, RTEs seek to unlock inter-agency coordination problems or operational bottlenecks and provide real time learning to the field.

The panel will focus on Inter Agency Real Time Evaluations in Mozambique, Philippines, Haiti and Pakistan for more information consult:

Keywords: Real Time Evaluation; Joint Evaluations
Participants and co-production II

Integrating evaluation into program design: a practical perspective on stakeholder involvement

M. Keene
United States Environmental Protection Agency

The United States Environmental Protection Agency (EPA) recognizes the value and the challenges of integrating evaluation into the design of new programs. Over the past two years, U.S. EPA’s Evaluation Support Division has collaborated with the Paint Product Stewardship Initiative (PPSI) to build evaluation into a pilot program for paint management in the State of Oregon. Leftover paint causes concerns for human health and the environment. It also incurs high costs on local and state governments who have generally borne the responsibility for managing the 65 million gallons of post-consumer (leftover) paint generated annually in the United States. The PPSI consists of participants from government agencies, manufacturers, trade associations, paint recyclers, and non-profits interested in the effect of the pilot program on consumer behaviors, paint collection and markets for recycled products. The PPSI wants to use evaluation to inform the design and roll-out of an industry-run, nationally coordinated paint management system. The Evaluation Committee began by articulating a conceptual model for the process of integrating evaluation into the pilot that focused on collaboration and efficiencies in cost, data collection, communication and evaluation utility. The committee used the model to develop evaluation questions, measures and a methodology; however, dynamic partnerships, politics and policies have introduced influential forces into the space between theory and practice, resulting in uncertainty and opportunity. The purpose of this evaluation is to provide the PPSI network, government agencies, industry representatives and citizens with evidence about the value of paint management. Waning enthusiasm amongst stakeholders and political pressures threaten the quality of evidence that the evaluation can generate. With an emphasis on data collection and access as well as use of results and learning, this paper describes the evaluation process and discusses limitations to stakeholder involvement in the design and systematic evaluation of a pilot program.

Keywords: Environment; Participation; Stakeholder; Program Design; Product Stewardship

Participants as Evaluators: Empowerment or Disenchantment?

R. Ellis1, E. Hoard2
1University of Chester, Social and Health Evaluation Unit, Chester, United Kingdom
2North Ontario School of Medicine, Programme Evaluation, Thunder Bay Ontario, Canada

While it is axiomatic that participants in a social programme should contribute to its evaluation, this usually stops short of participants actually undertaking the evaluation. This paper describes two attempts to involve participants in evaluating the programme they were receiving. One could be counted a success; the other a relative failure and lessons can be learned from each.

The first evaluation was of a Sure Start programme in the UK. Sure Start was a government initiative to give disadvantaged five year olds a better start in life. It was a condition of funding that each programme should be externally evaluated and we gained a contract to do that. The programme providers were keen that the parents of the children should be involved in the evaluation.

We approached this through involving the parents in surveying the views of participating and non-participating parents. Volunteer parents were given an opportunity to design, administer and analyze a questionnaire. In support they received a short research methods course. The questionnaire was designed in an idiom that spoke directly to the parents. A good response rate led to useful results, which were presented by the parents in partnership with the evaluators.

The second evaluation was of a Social Work training programme, also in the UK. The validating body for the programme insisted that they should be evaluated preferably with service user involvement. We aimed to involve service users of social work in the evaluation in a similar way to the Sure Start parents. However service users had an unrealistic expectation that they would be undertaking the evaluation supported by the evaluators and this approach proved problematic and ultimately unsuccessful.

This paper describes and evaluates these two experiences and relates lessons learned to participant evaluation principles and practice, politics and policy.

Keywords: Participation; Evaluation; Training; Politics; Policy
The active role of citizens in evaluation processes: methods, impact and prospects of civic evaluation in Italy

A. Tanese, A. Terzi
Cittadinanzattiva, Head Quarters, Roma, Italy

Civic evaluation has been called a comparative research-action, supported by declared and controllable methodologies and conducted by citizens in order to voice their point of view, issuing grounded judgements on relevant realities for the safeguarding of rights and the quality of life.

Over the past ten years Cittadinanzattiva has been the first no-profit organisation in Italy to promote and develop civic evaluation projects and methodologies in different PA sectors of intervention.

In the healthcare sector, in particular, through the Civic Audit methodology, in the decade 2001–2010 a joint group of citizens and operators in the sector has conducted a full cycle of evaluation in over 150 local health corporations and hospitals, which has had important cultural and organisational effects.

Other initiatives have been undertaken in the sectors of the environment, schools and local public services. Civic evaluation is currently being tried out, in collaboration with the Public Services Department, on the subject of urban quality in twelve municipalities of southern Italy.

Based on these experiences, the study firstly analyses the nature and characteristics of civic evaluation processes actually carried out, methodologies adopted and main results obtained.

In the second part the authors make reflections on the potential development and conditions for the success of civic evaluation as a tool for raising and reinforcing accountability processes in local public systems.

The basic idea is that a more active role for citizens appears to be essential for modernising evaluation systems already in place in various institutional ambits and sectors of the Public Administration. The authors believe that the introduction of the citizen’s “point of view” in evaluation processes in a systematic and technically valid manner may be an opportunity to rediscover the original meaning and basic values of public action.

Keywords: Civic evaluation, Healthcare
Evaluation policy: Policy and rural development II

Evaluation of EU-policy for rural development with focus on participation and collaboration of local actors

K. Pollermann, P. Raue, G. Schnaut
Johann Heinrich von Thünen-Institut, Institute of Rural Studies, Braunschweig, Germany

The Rural Development Programs (RDP) of the European Union support a wide range of activities in rural areas (main aspects are agriculture, environmental matters and quality of life). One part is LEADER, which is a bottom-up-oriented, participatory approach with a collaboration of local actors in a rural area (LEADER-Region), who are organised in a Local Action Group (LAG) as a kind of a public-private partnership.

One crucial objective of the evaluation in seven federal states in Germany is to detect the impacts of the participatory LEADER-approach: a core thesis of LEADER is, that including local actors in decision making improves the results of RDPs for rural development.

To explore the quality of the public-private cooperation in the context of LEADER several methods have been developed (case-studies, questionnaires, project data analysis). A core element to assess the functioning of the LAG and the benefits of participation is a written questionnaire, which was used in around 100 LEADER-Regions (altogether more than 1500 LAG-Members answered). In this context, to integrate the know-how of local actors, participatory methods were used for the development of the evaluation design and the interpretation of results.

In order to develop proposals for improvements in the future it is important to take into account the wishes of the regions as well as the objectives of EU-Policy. In the end of 2010 the so called Midterm-Evaluation will be finished. Some conclusions already can be presented about the questions: What evaluation methods can be used best to measure the impact of the participatory approach? What lessons for the further development of EU policy for rural areas can be drawn?

Keywords: Rural; Policy; Leader; Participation; Regional development

Approaches for assessing the impacts of EU Rural Development Programmes in the context of multiple intervening factors

European Evaluation Network for Rural Development, Helpdesk, Brussels, Belgium

A Working Paper on Approaches for assessing the impacts of the Rural Development Programmes in the context of multiple intervening factors has been published by the European Evaluation Network for Rural Development. Describing the possible methodologies to be applied for assessment of impacts within the context of the EU’s Common Monitoring and Evaluation Framework (CMEF) for the programming period 2007–2013, the document allows for a pragmatic handling of measure-specific impacts, and includes relevant explanation for the establishment and analysis of the counterfactual.

Rather than determining a common EU method for addressing these issues, the paper suggests a set of approaches to allow Member States to capitalize on the work they have already undertaken. From a methodological perspective, the three common socio-economic impact indicators of the CMEF (economic growth, employment creation, labour productivity) are more closely related than the four common environmental impact indicators (reversing biodiversity decline, maintenance of High Nature Value farming and forestry, improvement in water quality, contribution to combating climate change).

For all three socio-economic indicators the document recommends using propensity score matching (PSM) for constructing the control group. It also suggests calculating the difference in differences (DID), i.e. the combined comparison between two points in time (before-after) and between beneficiaries and non-beneficiaries. After having estimated the direct effects on programme beneficiaries, the indirect effects can be measured by appropriate methods to discount deadweight, substitution and displacement effects, and to take into account leverage and multiplier effects. The complexity of the environmental indicators requires focusing on various approaches, trends of baseline indicators and place-specific additional indicators. The methods for constructing control groups (including PSM and DID) can also potentially be applied for the environmental indicators. Data availability remains an important consideration overall.


Keywords: Assessment of impacts; Multiple intervening factors; Rural development programmes; Socio-economic; Environmental
Ongoing evaluation and evaluation systems in the cohesion and in the agriculture and rural development policies

J. Habuda
Hungarian Evaluation Society, Chair, Úröm, Hungary

The example of Hungary: Only a few significant evaluations in Hungary prior to 2009;
No evaluation system in place: No systematic approach in the assessment of evaluation needs, in the implementation of the evaluations and in the utilization of the evaluation results;

The evaluation system of cohesion policy programmes after 2009, its main system elements:
- Systematic approach – Evaluation plan for a 3 years period
- Purposeful market building
- High quality and ethic standards, publicity – supported by the Hungarian Evaluation Society
- Management assisted by a quality management process
- “Centrally driven”

Ongoing evaluation of the Rural Development Programme:
- Different regulatory environment
- Common indicators, “The CMEF”
- Stronger link between monitoring and evaluation
- Ongoing and systematic review of the indicators
- Mid-term evaluation in 2010
- Evaluability assessment of the New Hungary Rural Development Programme and the Strategy
- Evaluation Plan for 2009–2015 based on proposals of the Monitoring Committee
- Capacity building

Different funds, different approaches and common difficulties
- Result and impact indicators – quantification and data collection
- Closure of the 2004–2006 programming period – indicators
- Impact assessment, use of counterfactual
- Quality of evaluations
- Publicity of evaluations
- Utilization of evaluations

Possible synergies:
- Strengthen the ongoing nature
- Strengthen the system approach
- The use of ethic standards
- The use of quality standards
- New and improved impact assessment methodologies.

Keywords: Evaluation system; Ongoing; Cohesion policy; Rural and agricultural development; Hungary; Differences; Common difficulties; Synergies; New methodologies; Ethic standards; Quality standards; Publicity; Utilization; Evaluation society
Evaluation in research and higher education I

Network governance of the Finnish Strategic Centres for Science, Technology and Innovation
K. Haila, K. Lahteenmäki-Smith, P. Uusikylä, A. Eronen
Net Effect Ltd, Mikonkatu, FI-00101 Helsinki, Finland

Our study focuses on evaluating network governance of the innovation policy instrument “Finnish Strategic Centres for Science, Technology and Innovation” (hereafter referred to as SHOK, i.e. the Finnish acronym). The SHOKs are new public-private partnerships founded in the legal form of company (Ltd). The SHOKs aim to speed up innovation processes and to develop and apply new methods for cooperation between industry and the academia. There are currently six SHOKs in operation: within forestry, ICT, metal products and mechanical engineering, energy and the environment, built environment, as well as health and well-being.

How to evaluate such an instrument? In addition to governance networks, the SHOKs have features of both development and innovation networks. In our study, we have assessed the governance of the SHOK networks, as well as the opportunities and challenges they face. Furthermore, the methods for examining such networks and their structure are tested. For example, following questions are addressed: how does the SHOK network with its Finnish and international partners, what are the ways in which the SHOKs are networked within Finland, internationally, as well as between the SHOKs, and what are the governance solutions that have been developed?

Keywords: Network; Governance; Innovation; Science; Technology

Evaluation on EU Framework Programmes with reciprocal benefits: lessons from Germany and UK
A. Hsiao, A. L. Cheng
Chung-Hua Institution for Economic Research, Taiwan Division, Taipei, Taiwan

The paper aims to evaluate the structure, scope and strategies of the participation and networking in recent EU Framework Programmes. Despite EURAWATCH's (2009) doubts about effective interactions between the shaping of national priorities and the Framework Programme in all member states, however, this paper focuses on forming reciprocal benefits due to member states’ participation into FPs, including the FP support of developing R&D capacity and research quality at national level.

We identify the reciprocal benefits in two major domains: the interaction between Member States, EC and FPs; and the interaction among participants of FP projects. This represents the contribution to the success of science and technology at national and EU-levels, including the development of European and international partnerships.

Some arguments are emphasized on contrasting evidences of Germany and the UK. Germany and the UK have accounted for the highest share in the volume of participations, EC funding received and collaborative projects with international co-partners. While Germany’s overall performance is ahead of UK and other EU members, the UK has its leading position in specific programme, such as Fellowship and Networks programmes in FP6, with the second strongest research base in the world behind only the US. While German shares of participations in private industry, universities, and non-university research institutions, higher education institutions in UK have the highest participation level in FPs and the highest amount of funding. Lessons of UK and Germany in competitiveness from collaboration reciprocity are explored.

We identify mutual gains and complementary effects among UK versus Germany with collaborative co-partners; then we evaluate and compare the way R&D policy match with the priorities at national/EU-levels. Meanwhile, we establish evaluation criteria and index for programme analysis. The approach may contribute to understand the complementarity among research priorities of EU members in both the thematic priority-base and transnational collaboration-base.

Keywords: Framework Programme; Reciprocal Benefits; Research Collaboration; Programme Evaluation
The Ex-ante Evaluation of Taiwan S&T Programs

S. S. Li

STPI

The aim of a “good ex-ante evaluation” is to request the well-built framework and present the kind of logical thinking between the goals and indicators in programs’ proposals. For S&T programs’ reviewers, they can effectively review these proposals and get themselves desirable information from the proposals, so as to allocate the appropriate resources and budgets for S&T programs. And for S&T programs’ managers & staffs, they can logically follow the planned strategies to implement these activities, so as to present programs’ anticipated outputs and outcomes in the future. Taiwan governments, National Science Council, etc., make some standards on S&T programs’ proposals and agencies must follow these norms. However, these agencies are hard to present the logical structures of programs and emphasize on their anticipated outputs and outcomes. Therefore, the paper mainly discuss how these agencies should write a “good programs’ proposals” and use planning tools or elements under Taiwan S&T programs’ evaluation framework.

To emphasize on anticipated outputs and outcomes of Taiwan S&T programs, and to enhance the quality of programs’ proposals, the paper mainly combine some strategic planning elements, fishbone diagram, SWOT analysis, stakeholder analysis, goals or objectives analysis, and outcomes analysis, with United States Agency for International Development’s (USAID) Logical Framework Approach(LFA) responding to Taiwan’s S&T programs’ evaluation framework. Therefore, the paper includes several parts: (1) to realize the nature of S&T program planning elements; (2) to review the situation on Taiwan S&T programs’ evaluation framework; (3) to develop several planning elements responding Taiwan S&T program evaluation. By taking one of Taiwan S&T Programs as an example, the paper hopes to enhance the abilities of Taiwan program managers’ writing proposals and make programs have complete planning structures, so as to allocate the appropriate resources, even to become the reference of Taiwan S&T Programs.

Keywords: Strategic Planning; Ex-ante Evaluation; S&T Programs; Logical Framework Approach(LFA); Proposal Writing
Evaluation field building

Evaluation field building: models for strengthening capacity, knowledge, and structures

K. Hay 1, S. Sridharan 2, S. Da Silva 3, Z. Kun 4, F. Carden 5

1 IDRC, Evaluation Unit, Delhi, India
2 Centre for Research on Inner City Health, Department of Evaluation, Toronto, Canada
3 Sri Lankan Evaluation Association, President, Colombo, Sri Lanka
4 China Health Economics Institute, Peking University, Beijing, China
5 IDRC, Evaluation Unit, Ottawa, Canada

This panel describes active experiments in Asia, Europe and Latin America to build the field of evaluation. The panel starts from the understanding of a field as an area of specialized practice carried out by trained practitioners, rooted in practice, research, and theory based knowledge. The panel explores key elements of field building including expanding the knowledge base, supporting capacity building and emerging leaders, and fostering information exchange and collaboration. The panel lays out cases to demonstrate and articulate what this vision of field building looks like in practice and explores interconnections between the various elements. The session focuses on structures, systems, content, core competencies, and the dynamic linkages required in field building efforts. Field building is discussed through the exploration of a set of models including:

- An effort to generate knowledge on the impacts of health system reform in China while simultaneously building capacity, and structures of evaluation into the health system;
- Evaluation capacity building to redesign the health system from an inequity lens in Chile;
- An effort to build structures for long term training through a post-graduate diploma in evaluation in Universities in Bangladesh, India, Pakistan and Sri Lanka;
- An initiative to bring evaluators in South Asia together to strengthen practice, engage in peer learning and support, and share knowledge called the Community of Evaluators;
- An Evaluation Summer School organized for multiple years in Edinburgh, Scotland.

Differences in approaches and strategies between and across developing, transition, and developed economies will be discussed. Given the experimental and nascent nature of this work, critiques of the model and sharing of alternate models and experiments by the audience will be actively sought.

The panelists will include Katherine Hay, Kun Zhao, Soma de Silva and Sanjeev Sridharan. Fred Carden will be the chair of the panel.

Keywords: Evaluation field building; Capacity building; Evaluation curriculum; Health systems; Community of practice
How evaluation can influence positive social change

Missing the strategic wood for the tactical trees: the role of evaluation in revitalizing the activist role of oversight institutions

T. Beney, B. Williams
Feedback Research & Analytics, NA, Pretoria, South Africa

The democratic model of government is usually structured to ensure that the executive is held accountable by a legislature representing the interests of citizens. In modern democracies legislatures have become reliant on oversight institutions tasked with evaluation functions (e.g. National Audit Office in the United Kingdom, Netherlands Court of Audit) for information with which to assess and address the actions of the executive. Available examples of evaluations of the evaluators that perform an oversight function suggest that a common approach to such evaluations is emerging. This paper contends that the emerging approach, which focuses on the credibility and use of the outputs produced by these institutions, in other words their tactical contribution, is inadequate. If the evaluators are to be evaluated according to the emerging convention, they could perform favorably regardless of the level of functionality, stability or peril characterising the state of the democracy they are constituted to defend. However, it could be argued that against its broader function as an oversight body, sharing a joint mandate with sister institutions responsible for safeguarding democracy, the evaluators have not fulfilled their obligation if the democracy is under threat. This paper argues that in all democracies, but particularly in young democracies like South Africa, where the obligations of legislatures to citizens are contaminated by political expediencies and representatives are not necessarily competent, it is incumbent upon evaluators of the state to perceive the strategic landscape and assume an activist role in society. The evaluation of South Africa’s Public Service Commission, one of the constitutionally mandated bodies tasked with evaluating the delivery of services by the state, resulted in a set of recommendations that proposed it assume such an activist role. This evaluation is used to illustrate the argument.

Keywords: Democracy; Oversight; Activism

Organisational Evaluation as a Tool to Improve The Quality of Social Services in Portugal

P. Teixeira, P. Antunes
Logiframe – Consultoria e Formação Ltda, Evaluation Unit, Lisboa, Portugal

The Ministry of Labour and Social Security made in recent years a push for the quality improvement of the social services provided around Portugal by NGOs that are partly financed by that Ministry. We are talking of services that target the elderly and also special needs groups like the disabled or the homeless and about a vast number of services with very different natures, from kindergartens to resting homes. When trying to implement quality standards for these vast number of social services the Portuguese Social Security Institute was confronted with the incapacity of the majority of the NGOs to comply with most of them because they were lacking any kind of quality policies or management tools. After this diagnosis was made several programmes were launched regarding Quality with the objective of creating conditions for the real implementation of the quality standards for social responses.

This paper reflects on the practices of our team on one project implemented with sixty two (62) NGOs around Portugal and the role that organisational evaluation had as a tool for organisational change. We explore and detail the specific gains attained by these organisations and the publics for whom they work by the introduction of these processes. The effort made by these sixty two (62) on the last two years was rewarded with concrete and verifiable gains for there internal processes and dynamic, in the services they provided for their target groups and created a good basis for the effective implementation of the quality standards design by the Portuguese Social Security Institute. All this processes proved once more the importance and specific role that small evaluations can play on the efficient use of Public funds (national and european) and the effectiveness of National policies.

Keywords: Organisational Evaluation; Quality; Organisational Chance; Social Services; Social Security

Evaluation of policies subject to political and social debate: an evaluation of the management of intellectual property rights in Spain

J. M. Ivanov Iñiguez, C. Gomez de Marcos
Agencia Evaluación Políticas Públicas, Departamento Evaluación, Madrid, Spain

Certain policies are subject to significant political and social controversy. The evaluation of these policies can be a method to reduce social tension and to clarify positions in the social debate. In 2008 the management of intellectual property rights in Spain was holding a large and strong social and political debate. In this context, the Agency for Evaluation of Public Policies and Quality of Services was commissioned by the Spanish Cabinet Ministers to carry out an evaluation regarding to the intellectual property rights. Some key elements for the implementation of this evaluation were: the content of the assessment was developed in two phases with different levels of political intensity in order to achieve a sufficient level of confidence and prestige from evaluators in front of the stakeholders. The confidentiality requirements, the methodologies used and the scope of the evaluation. The problems faced by the assessment team and how to solve them make such assessments have special characteristics at the time of its development. The experiences gained by the assessment team can be a useful learning tool for other evaluators that address other issues or policies subject to an intense social and political debate.

Keywords: Evaluation; Social debate; Political debate
Methodological challenges III

How to evaluate “programmes” without programmes: Challenges in strategic level programme evaluation.

O. Oosi, M. Wennberg
Ramboll Management Consulting, Finland, Helsinki, Finland

In recent years there has been an increase in governmental programmes both at strategic and operative level. Former has become a dominant feature of public policy in Finland. These programmes are evaluated usually from the viewpoint of traditional “programme evaluation” which leads to challenges and contradictions with evaluation findings, uses of evaluation and stakeholder involvement.

Strategic level programmes are not programmes in the strict sense of the word. They are better described as governments’ “strategic priorities” or “horizontal development themes”. Typically, these programmes have high level political commitment, high level goals without specific measures and implementation mechanisms that resemble soft governance tools. These programmes may involve a funding scheme without any clear link to programme goals. First part of the paper describes this changing landscape of programme management.

Traditional programme evaluation concept, which has evolved in Finland mainly from structural funds evaluations and from the tradition of theory-based evaluation, is challenged: how to evaluate programme, which does not have any idea of its theory, programme organisation or “implementation mechanisms”, or even steering mechanism for its funding? Second part of this paper deals with these challenges and how they can be found in recent evaluation research and in four practical cases from Finland. It argues that traditional evaluation concept reveals the nature of some crucial problems in the concept of “programme” itself as in the management practice.

Third part of paper argues that this new situation calls for new ways how evaluation practitioners, programme management and evaluation users interact within evaluation process. It argues that when programmes are implemented as “strategic priorities” rather than full-scale programmes with programme organisation and detailed implementation plans, new kind of questions and approaches should be used in evaluation processes.

Paper uses four practical programme evaluation cases from different policy sectors from Finland.

Keywords: Strategic evaluation; Programme evaluation

Getting incentives right: do we need ex post cost-benefit analysis?

D. Sartori¹, M. Florio²
¹ CSIL – Centre for Industrial Studies, Development and Evaluation Unit, Milano, Italy
² Università di Milano, Dipartimento di Scienze Economiche Aziendali e Statistiche, Milano, Italy

This paper, presented at the Sixth European Conference on Evaluation of Cohesion Policy (Warsaw, 30 November-1 December 2009), discusses why there is a strong need of ex-post Cost-Benefit analysis and which conditions should be met for a proper ex-post exercise to be carried out in the framework of Cohesion Policy major projects. After an introduction about the objectives and instruments of the 2007–2013 EU Cohesion Policy, and in particular the legal framework for co-financing environmental and transport projects, the paper illustrates and discusses some methodological choices which have been made by the authors of the European Commission Cost Benefit Analysis Guide. It is showed that, without an ex-post Cost-Benefit analysis, the ex-ante exercise is also weakened as a decision-making tool. In particular, in the light of evidence from literature about the most common mistakes and pitfalls in ex-ante project appraisal, it is explained how systematic ex-post evaluation is important in particular linked to ex-ante incentives to reveal true information about the projects characteristics (especially on investment costs and demand forecast which are often respectively underestimated and overestimated due to an optimism bias) and ex-post performance assessment. The Commission has a unique role to play in this context, and recommendations are given about how to improve the use of Cost-Benefit analysis for investment decisions.

Keywords: Cost-benefit analysis; Ex-post evaluation; Cohesion policy
Kirkpatrick and beyond: A comprehensive methodology for influential training evaluations

J. Hense
Ludwig-Maximilians-Universität, Psychology, München, Germany

Background
The popular Kirkpatrick-“four levels” approach to training evaluation has a number of limitations when an evaluation is expected to systematically influence decision makers in public and private sectors. Kirkpatrick suggests to analyse four kinds of training outcomes: reactions, learning, behaviour and results. This presents an oversimplified view of training as it ignores individual and contextual influences. However, such data is essential when an evaluation’s goal is to inform training developers and decision makers on how to improve a training.

Objectives
The paper will analyse the weaknesses of the Kirkpatrick approach and its shortcomings in influencing political practice. Based on this analysis, it will outline a comprehensive methodology for influential training evaluations to overcome these problems. Conduct and results of this methodology will be exemplified by a case example of an evaluation study of trainings for Six Sigma, a methodology for organizational quality improvement.

Methods
The methodology systematically includes context and process analyses in addition to outcome data in a multi method approach to training evaluation. It includes interviews, structured observations of training sessions, pre-training, post-training, and transfer surveys.

Results
As demonstrated by the case example, the methodology was not only able to determine training impacts, but also several weaknesses which served as a source for recommendations for improving the training. An entirely outcome-focused approach such as the four levels would not have been able to provide this kind of formative feedback.

Conclusion
On a general evaluation theory and policy level, the paper can serve as an example of the limitations of pure impact evaluations. In many practical contexts they will not provide the information necessary to guide improvements of a program. As demonstrated by the suggested methodology for training evaluation, inclusion of context and process data is often indispensable for evaluation to sustainably influence practices and policies.

Keywords: Training evaluation; Formative evaluation; Evaluation use
Impact evaluation in the real world: dogma or good practice?

**Burt Perrin**: What is Rigour in “Rigorous” Impact Evaluation?
There are frequent calls for “rigorous” evaluation of development and other public sector initiatives. But too often “rigour” has been interpreted just in terms of research design. This presentation will identify other dimensions where rigour is important to meaningful impact evaluation, e.g.: rigour in matching the choice of design and methods to the situation and context; rigour in external as well as internal validity; rigour in implementation; in analysis, reporting, and in other characteristics necessary for meaningful and useful evaluation findings.

**Scott Bayley**: The Limitations of Econometric Methods for Impact Evaluation
Using econometric methods to undertake ‘rigorous’ evaluations of program impacts is a popular approach in development organizations. This presentation will review some of the more common shortcomings of this method and then propose evidentiary criteria for credible assessments of impact. By focusing on the fundamental principles for causal inference as identified by J S Mill evaluators are able to transcend debates about the gold standard and the merits of experimental vs. non-experimental approaches.

**Nicoletta Stame**: “Nothing works! What works? What can impact evaluation say?”
The debate on impact evaluation has so far mainly focused on issues of attribution of causality from the point of view of internal validity, even though limits have been recognized in the workability of “rigorous” methods that are unable to go beyond simple programs and stable situations. Yet, if what matters is providing policy-makers with evidence of what works or not, the main real world challenge is to be able to address problems of external validity, i.e. to generalize results across settings, contexts and policy domains, to build on diversity instead of uniformity. This also has significant implications for the conduct of syntheses, and requires a methodology that is “situational responsive”.

**Keywords**: Impact evaluation; Development evaluation
Evaluating good governance initiatives

Developing democracy: Measuring what works in advancing good governance

C. Elkins¹, M. Cohen²
¹ RTI International, Director M&E, Hillsborough NC, USA
² RTI International, M&E, Hillsborough NC, USA

Advancing good governance is a critical concern in international development, since foundational transparency and accountability have become prerequisites for assistance across most project sectors. The panelists have worked in international development across sectors for many years, and offer research designs and findings from current work in Guinea and Nigeria (plus other papers as available) that should interest evaluators across the spectrum.

Dr Elkins will present work in Guinea on a multisectoral project funded by USAID to promote good governance at the local level, toward more responsive services and participatory decisions. She explores the demand-driven approach being implemented in terms of innovative monitoring and evaluation approaches to capture the impact of facilitative assistance that works through local champions. (Separate paper abstract submitted: “Challenging champions: Evaluating good governance in motion”)

Dr Cohen will present work in Nigeria on a local governance project funded by USAID to promote good governance at the local level in two states, Bauchi and Sokoto. He explores the potential to use survey data collection to build our knowledge in the effective determinants of progress toward better governance and democratic participation through multiple channels. (Separate paper abstract submitted: “Establishing determinants of behavior relating to improved governance among constituents in Nigeria”)

We welcome other panelists who may be added to this session at the discretion of the Scientific Committee.

Keywords: International development; Governance; Survey research; Qualitative research; Mixed methods
Use of evaluation evidence for informed decision making under managing for performance results

S. Premakanthan
Symbiotic International Consulting Services (SICS), None, Ottawa Ontario, Canada

A generic definition of evidence in its broadest sense is a useful starting point. Evidence can be defined as “facts or testimony in support of a conclusion, statement or belief” and “something serving as proof” (Shorter Oxford English dictionary on historical principles). Such a definition is devoid of context and does not specify what counts as evidence, when, and for whom. In the evaluation context, we are producing and using evaluative evidence to determine the quality and worth by assessing the relevance, effectiveness and consequences of specific policies, programs and interventions. Evidence informed decision making is an endeavour in which there is informed, explicit, and judicious use of evidence derived from a variety of science and social research and evaluation methods. The round table session will discuss how the available evaluation evidence is used or not used by decision makers under managing for performance results, with examples from experience or drawn from research studies. The session will discuss the supply of quality, credible, timely evaluation evidence and some of the issues in gathering the needed evidence. The underlying assumption is that there is a demand for quality, credible and timely evaluation evidence by decision makers. More specifically, the session will explore the following: 1. how to support/encourage the use of evaluation evidence for policy, program and resource allocation decision making under managing for performance results (MfPR)? 2. how to generate the supply of quality, credible, timely evaluation evidence for policy, program and resource allocation decision making under MfPR? 3. What are the lessons learned on the impact of evaluation evidence on policy, program and resource allocation decision making?; 4. Is there sufficient Evaluation coverage?; and 5. What evidence do we have on the uses and drivers of effective evaluations.

Keywords: Use of evaluative evidence; Informed decision making; Evaluation and Program Management
Impact evaluation II

Evaluating for the public good: the experience of undertaking an organisational impact assessment of a charitable Foundation in the UK

K. Junge1, L. Thomas2, F. Sullivan2

1 The Tavistock Institute, Evaluation and Research, London, United Kingdom
2 The Health Foundation, Evaluation and Strategy, London, United Kingdom

The Health Foundation is an independent charitable foundation that works at every level of the UK healthcare system, funding complicated and complex improvement programmes that aim to develop the technical skills, leadership, capacity, knowledge, and the will for change, that are essential for real and lasting improvement.

Since 2006, the Foundation has been undertaking an internal assessment of the impact its investment in improvement interventions have had on the quality of healthcare in the UK. The purpose was to account for the organisation’s charitable spending on the one hand, and to inform the future strategic direction of the Health Foundation, and hence its allocation of charitable funds.

In line with the Foundation’s value of being grounded in evidence, the impact assessment, was designed to be scientifically rigorous. It involved:

- Using a theory of change and participative approach;
- Developing a measurement framework in consultation with internal and external stakeholders;
- Negotiating what was considered as evidence of impact, including commissioned evaluations, self-evaluation reports and programme monitoring data.
- Continually testing and refining the framework and approach, acknowledging where the limitations of the process were and seeking to address them for example, using case studies.

The final impact assessment paper was a collation of five thematic reports, arranged under the Foundation’s new articulation of how it works. Insights from the team leading the work included:

- Assessing the impact of an organisation within a complex system like the NHS is challenging;
- Faced with the same data, different internal stakeholders will arrive at different conclusions on the impact achieved. Arriving at a shared narrative is a negotiated process;
- A participative process can act as a catalyst for change, particularly for a relatively young organisation working within the emerging field of quality improvement and where internal buy-in is strong.

Mix methods impact evaluation design for the Mesoamerican health initiative: quantitative component

J. P. Gutiérrez Reyes1, M. Romero Martínez1, M. M. Téllez Rojo1, M. d. P. Torres Pereda2

1 Instituto Nacional de Salud Pública, CIEE-INSP, Cuernavaca, Mexico
2 Instituto Nacional de Salud Pública, CISS-CIEE, Cuernavaca, Mexico

Background

The Mesoamerican Health Initiative (MHI) is a joint regional effort which seeks to reduce health gaps in Mesoamerican populations. To this end, MHI has established four thematic axes according to the most relevant health conditions in Mesoamerica: nutrition, maternal and child health (MCH), vaccination and vector-borne diseases (malaria and dengue fever). The MHI is an 8 country regional intervention which seeks to deliver simultaneously a package of proven interventions and strategies integrated synergistically and targeted in vulnerable populations: rural indigenous and urban low-income groups.

Objectives

To present the design of a mix methods impact evaluation for the MHI with emphasis on the quantitative component.

Methods

A regional group of experts was designated with the objective of design the impact evaluation plan for MHI. 3 regional meetings were held to set this participative evaluation design.

Results

Randomize controlled experiment with a baseline study and two follow-up studies (2 and 4 years) using a sequential mixed method approach will be carried. The evaluation comprises of a main panel stratified by rural and urban zones where information about nutrition, MCH and vaccination; whereas vector-borne diseases will be evaluated in supplemented panels. Localities, households and individuals will be selected randomly, visiting primary health care centers too. Units will be sampled by pairs of clusters with random assignment of treatments. Sample size of 1,000 households in 30 localities selected by cluster by group (2,000 households and 60 localities total) would be adequate for countries with low burden of malaria and dengue. Through surveys and biomarkers, differences in differences analysis will be carried out.

Conclusions

However evaluating 4 thematic areas and an anticipated heterogeneous implementation encompassing activities in 8 countries with a design that must be covered under one single evaluation was a challenge, a feasible and complex design was created in a participative way.

Keywords: Impact evaluation; Mix methods; Mesoamerica; Health interventions; Quantitative component
Mix methods impact evaluation design for the Mesoamerican health initiative: qualitative component


1 Instituto Nacional de Salud Publica, CISS/CIEE, Cuernavaca, Mexico
2 Instituto Nacional de Salud Publica, CRISP-INSP, Cuernavaca, Mexico
3 Instituto Nacional de Salud Publica, CIEE-INSP, Cuernavaca, Mexico

Background
The Mesoamerican Health Initiative (MHI) is a joint regional effort which seeks to reduce health gaps in Mesoamerican populations. To this end, MHI has established four thematic axes according to the most relevant health conditions in Mesoamerica: nutrition, maternal and child health (MCH), vaccination and vector-borne diseases (malaria and dengue fever). The MHI is an 8 country regional intervention which seeks to deliver simultaneously a package of proven interventions and strategies integrated synergistically and targeted in vulnerable populations: indigenous and low-income groups in rural and marginalized urban zones. Facing this challenge, rigorous evaluation was needed.

Objectives
To present the design of a mix methods impact evaluation for the MHI with emphasis on the qualitative component.

Methods
A regional group of experts was designated with the objective of design the impact evaluation plan for MHI. 3 regional meetings were held to set this participative evaluation design.

Results
Randomize controlled experiment with a baseline study and two follow-up studies (2 and 4 years) using a sequential mixed method approach will be carried. The evaluation comprises of a main panel stratified by rural and urban zones where information about nutrition, MCH and vaccination; whereas vector-borne diseases will be evaluated in supplemented panels. Based on a constructivist perspective, meanings and practices will be analyzed to achieve a summative evaluation. Purposeful sub-sample for maximum variation will be recruited based on 3 criteria: localities, households and individual characteristics. 21 localities for main panel, 14 for the supplemented ones and 16 control localities will be studied using semi structured interviews, discussion groups and non participative observation.

Conclusions
However evaluating 4 thematic areas and an anticipated heterogeneous implementation encompassing activities in 8 countries with a design that must be covered under one single evaluation was a challenge, a feasible and complex design was created in a participative way.

Keywords: Impact evaluation; Mix methods; Mesoamerica; Health interventions; Qualitative component
Country based and country led evaluation

Country level Monitoring and Evaluation: A South African Experience

J. Mathe¹, T. Masangu², N. Gcwabe³

¹Director: Programme Evaluation, Public Service Commission, South Africa – JabuM@opsc.gov.za
²Deputy Director: Programme Evaluation, Public Service Commission, South Africa – ThokoM@opsc.gov.za
³Deputy Director: Programme Evaluation, Public Service Commission, South Africa – NoqoboG@opsc.gov.za

The South African Government has, over the past fifteen years, recognized the challenges arising from having to transform the state into a developmental one while simultaneously trying to redress the legacy of past social imbalances. To ensure that tangible results are achieved, there has been an increasing recognition of the urgent need for greater effectiveness in the way that government monitors and evaluates its policies, projects and programmes.

This paper discusses the various initiatives taken by the South African Government to institutionalize M&E in Government. This includes the establishment of a new Ministry of Performance Monitoring and Evaluation in the Presidency in 2009, the Government Programme of Action Reports, the publication of the policy document, entitled Improving Government Performance: Our Approach, the publication of the overarching Policy Framework for the Government-wide Monitoring and Evaluation System and the publication of reviews which summarise data on trends to track progress towards the implementation of government programmes.

Keywords: Monitoring and Evaluation; South Africa

New Approach for Country Program Evaluation (CPE): recent CIDA’s Experience

R. Beaulieu
Canadian International Development Agency (CIDA)

Objective

Many members of the Development Assistance Committee (DAC) are requested to undertake evaluation of their programs of development cooperation in partner-developing countries. In that context, the Canadian International Development Agency has developed a standardized approach focussing both on results and the assessment of the performance of development cooperation programs in a given country or region.

Background

The approach is consistent with the Paris Declaration principles (Ownership, Alignment, Harmonization, Management for Results and Mutual Accountability) and is based on key criteria accepted by the DAC Development Evaluation Network including: i) Relevance, ii) Effectiveness, iii) Sustainability, iv) Cross-cutting issues, v) Coherence vi) Efficiency, vii) Management Principles, viii) and Performance Management.

Methodology and Approach

Moreover, the approach provides a methodology allowing to compare the different delivery mechanisms to potentially find out what work better. This component of the methodology is quite innovative, given the current state of the discussion regarding program-based approaches compared to the traditional project approach. The study provides an analysis based on a sample of 118 projects reviewed in five countries: Bangladesh, Ethiopia, Malawi, Mozambique and Vietnam.

Results/Conclusions

The overall conclusion of the study indicates that the various delivery mechanisms utilized by many actors of the international community all have strengths and weaknesses. It is too early in our view to declare that the project approach is outdated or that the program-based approach including general and sectoral budget support is the only delivery mechanism that deserves to be part of the tool box of development cooperation agencies.

The approach was presented at the recent Canadian Evaluation Society Conference in Victoria, B.C. Canada and has been shared with different members of the DAC Evaluation Network at a recent meeting in Brussels.

Keywords: Country Program Evaluation; Developing Countries; Performance Assessment; Comparing different delivery systems
Development of evaluation in Ukraine

I. Kravchuk
National Academy of Public Administration under the Office of the President of Ukraine

Background
Ukraine has quite fragmented evaluation system in public sector. Some evaluation methods are applied in the legislative and budgetary process, strategic planning. Brave “pilot projects” on application of evaluation philosophy in European integration were drivers of cultural change in public administration. Evaluation methodologies are used in international projects and accepted by Ukrainian partners in non-governmental sector. But it is not enough for wide dissemination of result-oriented management.

Objectives
Evaluation of meeting strategic objectives at the political level and performance of public institutions activities supports transparency and accountability, effectiveness and efficiency of governance. Therefore important questions for Ukrainian public sector are:
• How can we create demand for ex-ante and ex-post evaluation of public policies?
• What are the main challenges and obstacles in establishment of comprehensive evaluation system?
• How can we integrate evaluation to normal strategic planning and management process?
• How institutional capacities and evaluation culture can be developed?
• What foreign experience can be the most applicable for Ukraine?

Methods
Relevant Ukrainian legislation and practice is analyzed, civil servants and practitioners are interviewed in the framework of this research. Recommendations will be presented for expert community.

Results
Expected practical results of the Policy Paper are:
• Current situation in Ukraine is analyzed.
• Possible gaps and problematic areas are identified.
• Lessons from pilot project of establishment of M&E system of EU-Ukraine arrangements are learned and recommendations for other sectors are elaborated.
• Proposals for improvement of evaluation system in Ukraine are developed.

Application to practice
Proposals for improvement of evaluation system in Ukraine should be translated into concrete legislative initiatives and institutional solutions. These ideas and recommendations can be basis of international development projects on administrative reform and support to implementation of EU-Ukraine Association Agreement. Final recommendations will be submitted to Ukrainian authorities.

Keywords: Evaluation in public sector; Evaluation capacities; Governance
Thinking learning and doing evaluation: A report on local capacity building

Thinking, Learning, and Doing Evaluation: A Report on Local Capacity Building


1 University of Boras, FoU Sjuhärad Välfärd, Boras, Sweden
2 University of Illinois, pedagogic, Urbana-Champaign, USA
3 Hedmark Univ College Norway Karlstad University, FoU Välfärd Värmland, Elverum Norway Karlstad, Sweden
4 Mälardalen University, Evaluation academy, Eskilstund, Sweden

The Swedish Evaluation Verkstad Practice (EVP) is a means to enhance the capacity of human service workers and their organizations to conduct useful and defensible evaluations. The EVP engages a group of 10–15 welfare, health, or education professionals in an adult learning process, facilitated by university and R&D center evaluation experts. The professionals come together to learn about evaluation by actually conducting an evaluation. All participants bring along an assignment from their organizations to conduct an evaluation of a particular program or initiative. The group meets for a full day about once a month for a period of about 12 months. This session highlights the underlying logic or program theory of the EVP and present recent data collected on EVP experiences and outcomes from apr. 100 participants in EVP and their managers. Jennifer Green, chair of the panel, introduces the thinking behind EVP.

First Presentation Title: An Overview of Evaluation Verkstad Practice. This presentation will briefly describe the overall structure and rationale of the EVP. EVP constitutes an activity that accomplishes the combined purpose of conducting evaluations and developing competence to conduct evaluations, with the support of R&D units and university-based evaluators.

Second Presentation Title: Who comes to Evaluation Verkstad Practice? “All are welcome” is the overarching message, but at the same time the participants need to have time to participate, which requires negotiations with other actors in an organization.

Third Presentation Title: The Program Theory of EVP – Further Steps. This paper explores the designed and implemented program theory of EVP, compared to the experienced one.

Fourth Presentation Title: The result of EVP. This paper concludes the session by identifying valued outcomes of the EVP: completed evaluations for human service organizations, local learning about evaluation, and contributions to evaluation capacity building.

Keywords: Internal evaluation; Learning by doing; Capacity building; Program theory; Welfare organisations
Sectoral policy: Gender

Demographic change and gender equality dimensions within ERDF Programms: an ex post evaluation

F. Pesce¹, M. Samek Lodovici²

¹ IRS – Istituto per la Ricerca Sociale, Training and Labour Policies, Bologna, Italy
² IRS – Istituto per la Ricerca Sociale, Labour and Industrial Market Policies, Milano, Italy

Background
This paper presents the results of an ex post evaluation of “Effectiveness of the Cohesion Policy: Gender Equality and Demographic Change” conducted by a Consortium formed by IRS and CSIL for the European Commission, DG Regional Policy. This report was part of the general ex post evaluation of Cohesion Policy Programmes 2000–2006 co-financed by the ERDF (Objective 1 and 2).

Objectives
Main aim of the ex post evaluation was to provide an assessment of the extent to which the ERDF supported regions in adapting to demographic changes and fostering gender equality, as well as a better understanding of what ERDF interventions can deliver in this respect during future planning periods. On this basis, the paper will present the context of the evaluation, key evaluations questions and research methodology as well as main findings, lessons learnt and policy implications for future.

Methods
The study was based on:
– an extensive review of the literature on gender equality and demographic trends in order to derive the analytical framework and formulate working hypothesis;
– the selection and analysis of those ERDF fields of interventions that were most likely to influence gender equality and demographic change either directly or indirectly. Research hypotheses were formulated regarding the extent and typology of effects that could be observed;
– 12 regional case studies for field analysys to test the working hypotheses;
– 6 mini case studies on selected ERDF projects implemented in regions considered as examples of good practices.

Results and Conclusions
This evaluation proved that ERDF may indeed support regions in adapting to demographic change and fostering gender equality identifying specific main ERDF interventions and effective results through the case studies. Nevertheless, many are the efforts that should be increased for future planning to really internalised in ERDF programmes the gender mainstreaming approach and to cope with demographic changes. Several policy implications are presented in order to achieve these aims.

Keywords: Ex post evaluation; Gender equality; Demographic changes; ERDF programmes

Evaluating the European Social Fund in a gender perspective: methodological issues

S. Sansonetti, P. Casavola, A. Rosselli
Fondazione Giacomo Brodolini, Gender, Roma, Italy

Since September 2009 and until August 2010 Fondazione G. Brodolini has participated to an evaluation study on the European Social Fund’s support to the gender equality.

Strategies for gender equality both gender mainstreaming and specific actions have been considered focusing on the programming process and the preliminary stages of implementation as it emerges from document analysis and interviews with informants in of ESF interventions for 2007–13 in all European Member States.

The evaluation considered the Logics of ESF Intervention in each country in a gender perspective identifying the country’s approach in the use of ESF funds for gender policy. To obtain meaningful results from the country comparison a methodology for clustering countries has been developed. It has been structured according to a two-step approach.

• The first step has identified countries in similar conditions before the beginning of the Programming Cycle (2007–2013). To this aim reference has been made to a Gender Equality Index (developed by Plantenga et al.) and countries have been grouped accordingly.

• The second step has made use of an indicator constructed specifically for this exercise, the ‘ESF Policy effort indicator for gender equality’ in each country as it has emerged by the study itself (from the ESF programming and first implementation phase). This index has two values: above or below a discriminating value. It was obtained as a weighted average of two sub-indicators on the ESF resources allocated to gender positive actions and to the Gender equality institutional organization as it emerged from the ESF logic of gender intervention in each country.

This paper will discuss the methodological aspects of the study in general and more specifically the clustering procedure adopted and the results of the clustering of countries that will finally emerge from the study results.

Keywords: Gender; ESF 2007–2013; Methodological issues
“Evaluation responsive to gender and human rights: from policy commitments into action”

B. Sanz Luque, S. Reddy
UNIFEM, Evaluation Unit, New York, USA

The United Nations core mandate is grounded on the advancement of social justice and equality as reflected in Universal Declaration of Human Rights and all subsequent international human rights instruments that address different forms of discrimination and inequality through the commitments made by Member States to respect, protect and fulfill the rights they contain.

All UN organizations are also mandated to promote human rights and work towards the elimination of inequality, thus having to evaluate whether their programming has successfully contributed to the translation of these commitments into action and results. However, the knowledge-base and experience for substantively evaluating the gender equality and human rights of UN programming is limited and requires the evolution of evaluation thinking, approaches and methodologies to fully address these complex issues.

The United Nations Evaluation Group (UNEG) has been making progress in this regard through the adoption of the evaluation Norms and Standards with clear references to the integration of these dimensions, and more recently, with the development of specific guidance to assist evaluators in substantively integrating gender equality and human rights in all stages of the evaluation process.

United Nations Development Fund for Women (UNIFEM), a member UNEG member, has a specific mandate to support the advancement of gender equality and women’s rights, where evaluation is becoming a cornerstone of understanding what constitutes effective programming in UNIFEM as it works towards its goals. UNIFEM also plays a role in advocating for gender equality and human rights responsive evaluation within the broader UN system.

However, integrating gender equality and human rights in evaluation can be challenging. It requires a conducive environment, resources, capacities, commitments and knowledge to transform the way in which evaluations are planned, designed, conducted and used.

This paper presentation will address the evolution of gender equality and human rights responsive evaluation in the experience of UNIFEM, share progress made to date and highlight challenges identified and areas that require further thinking and exploration.

Keywords: Gender; Evaluation; UN experience
Evaluation and transition economies

Institutionalising the EU structural funds evaluation in Lithuania
D. Burakiene
Ministry of Finance of the Republic of Lithuania, EU Cohesion Policy and Structural Assistance Coordination Department, Vilnius, Lithuania

In the pre-accession period regulatory requirements of EU membership forced the administration of Lithuania and other Eastern European countries to adopt various instruments of evidence-based management, including evaluation of the EU Structural Funds. Lithuania is considered to be a regional front-runner in the area of public management reforms. According to the World Bank, the top-down approach of the Lithuanian administration produced quick results in the area of strategic planning. Another success story, which is the focus of the proposed presentation, is the well-institutionalised system of the EU Structural Funds evaluation.

There are two main stones of the EU Structural Funds evaluation system in Lithuania. First – the main initiatives of the evaluation activity are concentrated in the Ministry of Finance, which prepares annual evaluation plans, sets up a number of evaluation standards and carries out capacity building activities (training, methodological guides or publicity events). Second – decentralized nature of the system, when all actions, proposed evaluation themes and implementation problems are discussed in the Inter-institutional Evaluation Co-ordination Group. Such balanced approach to the management of evaluation allowed a quick roll-out of the new evaluation system. This contributed to strengthening the evaluation culture in public service.

At this stage Lithuania’s administration is facing new important challenges in the area of evaluation. Following an assessment of Lithuania’s evaluation system analysis, there is a key challenge to promote awareness of the benefits of evaluation (especially at the politicians level) and stimulate the needs and evaluation initiatives in the line ministries in order to realizing the potential of decentralized system of evaluation in Lithuania. Second, national budget programmes should also become a subject of evaluation. Therefore, the evaluation system should be extended and based on the successful experience of the EU Structural Funds evaluation.

Keywords: Lithuania; EU Structural Funds evaluation; Evaluation system; Evidence-based management

Current developments in the evaluation practice in Romania – from Phare to Structural Instruments
M. Nica1, C. M. Profiroiu1, T. Andrei2
1 Academy of Economic Studies, Management, Bucharest, Romania
2 Academy of Economic Studies, Statistics and Econometrics, Bucharest, Romania

The evaluation of the European Union funded programs developed rapidly in recent years, becoming more rigorous, systematic and sophisticated over the successive programming periods. Introduced in the pre-accession context in 1996, Phare Interim Evaluation developed into an important practice in the European Union. From 2000 the European Commission stipulated that all aspects of its interventions should be evaluated, therefore the requirement that Phare programs to be evaluated was embodied in the Phare Guidelines 2000–2006.

Prior to its accession to European Union, in January 2007, Phare Interim Evaluations in Romania were organized by the European Commission services in Brussels. After the accession and the introduction of the Extended Decentralized Implementation System (EDIS) this responsibility was transferred to the Romanian authorities.

The paper is trying to examine how the experience gained throughout the evaluation of the pre-accession funds generated knowledge for learning, therefore contributing to more efficient evaluation of the Structural Funds and public policy in Romania. Currently, while the last Interim Evaluation of Phare and Transition Facility from Romania is carried out, the first round of evaluations on Operational Programs are unfolding.

A key question is whether the institutional setting, administrative environment and legislative framework are conducive to a sustainable evaluation practice and whether the function of evaluation, as tool for policy learning and management, is widely understood. Moreover, the use of the evaluation is scrutinized with a view to growing constraints on government spending due to the global financial crisis.

Keywords: Evaluation; Public policy; Institutional capacity; Impact
Evaluation – an exception or a rule of thumb in Estonian innovation policy

K. Mannik¹, R. Rannala¹, A. Reid²
¹ Technopolis Group, Tallinn Office, Tallinn, Estonia
² Technopolis Group, Brussels Office, Brussels, Belgium

Based on INNO-Policy Trendchart country comparisons evaluation as an input in Estonian innovation policy is more usual practice compared to other Central and Eastern Europe countries. However, this phenomenon in policy-making cannot be still considered as a rule of thumb in Estonia. At programme level the evaluation requirement is not written into legal acts of programmes and therefore cannot be considered compulsory (like auditing) in policymaking.

The task of the current paper is to present the existing evaluation track record (including aim, methods, results, policy implications) of Estonian innovation policy. The practice goes back to the beginning of 2000s. The first evaluation of Estonian Innovation System was carried out by Hernesneimi, H. (ETLA) in 2000, which was commissioned by the Ministry of Economic Affairs. It has followed by three other wide-scale innovation system and/or policy evaluations during ten years. Complementary to it, ex-ante and mid-term evaluations of certain programmes or funding sources (European Union Structural Funds) have been executed for policy planning. Technopolis Group has given external support to Estonian innovation policymakers for the last decade.

Based on the analysis of evaluation practice in Estonia the paper also includes relevant international practice. As a result the paper provides a map of applied evaluation methods and methodologies in Estonian innovation policy. This is accompanied by recommendations to policymakers with the focus to set up regular policy evaluation in Estonia. More generally, the quality and contents of particular type of evaluation varies across programmes and policies which needs to be also tackled by policymakers in Estonia.

Keywords: Innovation policy; Evaluation practice in Estonia
Enabling or Constraining? A Multiple Case Study of the Impact of Accreditation on Canadian University Business Schools

C. J. Elliott, S. C. Goh
University of Ottawa, Telfer School of Management, Ottawa, Canada

Background
The question of AACSB® accreditation and its impact on Canadian university business schools has been studied empirically in only a limited manner in recent years. The purpose of this study was to explore the potential consequences – both enabling and constraining – of accreditation as perceived by administrators and faculty members.

Methods
A qualitative, multiple case study approach was employed. A purposive sample of four Canadian business schools was selected; and data were collected from multiple sources. The data were analyzed using NVivo7 and individual case profiles were prepared. These profiles were validated and a cross case analysis was performed.

Results
Informed by Adler and Borys’s (1996)® notion of bureaucracy and its enabling or coercive effects, the results of this study show that accreditation was not perceived to be overly constraining, in a negative way. Overall, AACSB accreditation was seen to be slightly enabling or neutral in its effect. While accreditation was viewed as a possible mechanism for increasing workplace formalization, this was not borne out by the evidence. The participants at three business schools felt that formalization was already quite high and they did not believe that the degree of formalization had changed with accreditation. However, at the highest impact school, accreditation was seen to increase formalization; and its effects were seen to be primarily enabling.

Application
This research has implications for practice at institutes of higher education, particularly university business schools. For university Deans and Directors, this research can provide some key elements to consider when deciding whether to embark upon an accreditation journey. It also provides insight into the contextual factors which contribute towards enabling effects (e.g., continuous improvement, organizational learning) and how to facilitate sustainability and institutionalization of these practices.

Direct public funding for research, (technological) development and innovation in Austria – selected findings from a meta-evaluation on a systems level

S. Mayer, I. Fischl, J. Streicher
Austrian Institute for SME Research, Evaluation Innovation and Technology, Vienna, Austria

In 2008–2009, the so called “Systems Evaluation” (Evaluation of Government Funding in RTDI from a Systems Perspective in Austria) was conducted by a consortium of Austrian Institute of Economic Research (Lead), the Austrian Institute for SME Research, Prognos AG (Berlin), and Convelop, including various external experts. It was commissioned by the Federal Ministry of Transport, Innovation and Technology und the Federal Ministry of Economy, Family and Youth.

The paper will present the main findings from the analysis of the system of direct public RTDI funding in Austria on the federal level; and will integrate the users’ perspective, focusing on the strategic behaviour of companies and research institutions as potential beneficiaries when approaching the “funding system”.

Various instruments are needed to properly address the different requirements and target groups of RTDI funding. Different logics of funding and the attempt to develop specific offers for each objective and target group lead to a multitude of programmes on the one hand; on the other hand the target groups face a system that is characterised by diversity and that offers incentives which partly compete with each other and seem little coordinated so far. At the same time the users of RTDI funding schemes react less specifically to the diverse offers of direct RTDI funding than was to be expected.

The findings suggest that a broader concept of RTDI policy and a better coordination of different funding schemes addressing different topics on the level of the decision makers might enhance the effects of public funding, while the necessity of highly specific programming can be disputed. As a necessary prerequisite, the suggested system would need to be grounded on a sound basis of overarching monitoring and impact indicators; this is gaining even more importance in the light of the new budget law in Austria.

Keywords: RTDI funding; Meta evaluation; Impact of evaluation; Context of evaluation
A comprehensive evaluation on Brain Korea 21 Phase II program

S. J. Kim, S. C. Byeon
Korea Institute of S&T Evaluation and Planning, HRST Policy Division, Seoul, South-Korea

In the late 1990s, Korean government in response to concern over the relatively low standing of the nation’s universities and researchers, launched the Brain Korea 21 Program (BK21). BK21 seeks to nurture globally competitive research universities and graduate programs and to breed high-quality research manpower in Korea. It provides fellowship funding to graduate students, post doctoral fellows, and contract-based research professors who belong to research groups at top universities. In Phase II, which began in 2006 and is scheduled to run through 2012, BK21 allocates about US $ 260 million a year. Phase II emphasizes department-level excellence and the university-industry link.

BK21 has attracted a great deal of attention, in part because of its strategy of concentration that results in relatively large awards and because of the prestige it confers on recipients. Given the high profile of the program, there is great interest in the program and in determining its effects on universities and human resource development.

In this study, we develop an evaluation model to assess the net effect of the BK21 Phase II program. The evaluation model includes a logic model, structural equation, regression model. We propose a logic model to identify the goals and missions, inputs, incentive, activities, outputs, and outcomes of the program and to explore the logical linkages among these elements. We develop the logic model to understand underlying policy intentions and program dynamics. The structural equation aims to assess the impact on BK21 program. The regression model aims to assess the net effect of the program in terms of SCI papers. Somi Seong, et al(2008) analyzed the net effects of BK21 program by difference-in-differences model. But, there is selection bias in this model. We try to solve this problem using propensity scores matching. It is difficult to collect data but we expect to measure more exact net effects of BK21 program.

Keywords: Brain Korea 21; Evaluation; Net effects; Propensity score matching; Structural equation
The interplay between politics and evaluation: options stakeholders, priorities and decision

Policy evaluation by government inquiry: Whither the emerging community safety approach to natural hazards in Australia?

G. Elsworthy 1, A. Rhodes 2

1 Deakin University, School of Health and Social Development, Melbourne, Australia
2 Royal Melbourne Institute of Technology, School of Global Studies Social Science and Planning, Melbourne, Australia

Background
Tropical cyclones, floods, severe storms and bushfire (wildfire) are significant sources of loss of life and property in Australia. Several recent events have taken a heavy toll. In the absence of clearly integrated and communicated policy, major sources of guidance to governments and their agencies have been the diverse recommendations and discussion in a series of government inquiries commissioned following significant hazard events, particularly bushfire.

Method and Results
Detailed content analysis of six recent inquiry reports supported the development of a comprehensive values-based framework for the evaluation of community engagement and education activities for natural hazards. The reports were characterised by a transformation in thinking from agency response to household and community responsibility and preparedness with an emphasis on shared (community/agency) responsibility, community participation in planning and program design, and partnership (the ‘community safety approach’). A theory-generating (‘realist’) synthesis of evaluation studies of fourteen Australian community engagement and education programs supported the approach, yielding an overarching program-theory model and principles of good practice.

However following a catastrophic event in Victoria (bushfires in February 2009) a Royal Commission with strong judicial powers has been established with very broad terms of reference. In contrast to the previous inquiries, the structure and processes of the Commission have had a predominant focus on organisational and operational failure, agency responsibility and system effectiveness.

Conclusion
The recommendations of the Commission’s Interim Report focus on a range of technological solutions and improvements to various systems and processes but contain negligible discussion of community engagement and education policy and practice, signalling a shift to a role for agencies as ‘protectors’ of the community as opposed to the partnership approach that had emerged over the past two decades. Selected results of the research project outlined above will be discussed in the context of this emerging shift in thinking.

Keywords: Politics; Policy evaluation; Natural hazards; Community engagement; Community education

Metaevaluation of evaluations on sustainability policies performed by the spanish national agency of evaluation of public policies

F. Prieto 1, J.A. Errejon 1, M. Macia 1, A. M. Lopez 1, J. M. Diaz 1

1 Agencia Estatal de Evaluación de las Políticas Públicas y la Calidad de los S, Evaluation, Madrid, Spain
2 Agencia Estatal de Evaluación de las Políticas Públicas y la Calidad de los S, Quality, Madrid, Spain

Metaevaluation is an integrated approach that requires a high degree of abstraction and coherence. The proposed work attempts to analyse from this perspective four evaluations of public policies performed by the Spanish National Agency of Evaluation of Public Policies related to sustainability issues: analysis of the registry of greenhouse effect gasses, influence of socioeconomice activities in natural park areas, road safety plan and management of watershed confederations.

This four evaluations include aspects related to sustainability, such as greenhouse gasses emissions, preserving nature and landscapes, continental water policy and public safety, which is related to risk management, infrastructures and atmospheric pollution.

Following Scriven (1991, 2007) and his Key Evaluation Checklist, we will perform an approximation to metaevaluation, adapting it to the context and the items of the performed evaluations. In particular the following aspects have been analysed:

1. framework and context of evaluation
2. analysis of the mandate
3. methodologies used in each evaluation
4. singularities of methodologies used in each evaluation
5. short description of evaluations
6. identification of stakeholders and public
7. analysis of main conclusions and recommendations

Keywords: Metaevaluation; Sustainability; Public policies; Spain; Scenarios
Democratic knowledge production and learning in the evaluation for non-governmental organizations in development

T. Kontinen¹, K. Holma²
¹University of Helsinki, Institute of Development Studies, Helsinki, Finland
²University of Helsinki, Research Unit for Social and Comparative Studies in Education, Helsinki, Finland

Non-governmental organisations (NGOs) occupy an increasingly important role in the international development aid. With their growing significance, the demands for showing effects and impacts by means of rigorous, often experimentalist, evaluation practices have been uttered; not least by the international donors that channel a remarkable portion of their development aid budget through the NGOs. In addition to the self-evident accountability purposes, the NGOs are often committed to learning in evaluation. Moreover, a great number of NGOs subscribe to the idea of democratic practices in North-South relationships. Thus, in terms of evaluation methodologies the experimentalist evaluation may conflict with NGO’s own ideology. In this paper we first argue that the democratic ways of knowledge production do not contradict with the aim of receiving adequate knowledge of the evaluated projects. On the contrary, we suggest that the democratic knowledge production improves the epistemic status of knowledge by integrating various expert and practitioner perspectives considering the subject matter. Second, we propose that learning in evaluation should be conceptualised as a knowledge production that provides new tools and ideas for the next steps for action for those whose living circumstances are at stake. In our attempt to conceptually connect democracy, knowledge production, and learning we consider the Deweyan account of community of inquiry as a context and means of evaluation that resonates with NGOs own search for evaluation designs alternative to experiments. Further, in Deweyan spirit, we define the concept of learning as the transformation of practices. In addition, we critically examine the possibilities of such knowledge production in this particular context characterized by accountability demands and unequal North-South relationships. We provide illustrative examples on evaluation discussion in Finnish NGOs engaged in development.

Keywords: Community of inquiry; Non-governmental development organisations; Learning; Democratic knowledge production
Participants and co-production III

Evaluation and participation of stakeholders: An example of a participative methodology regarding the Operational Programme against discrimination in Spain

V. S. Esteban
Red2red Consultores Sl

The paper we are presenting describes the design and the methodology of a participative evaluation undertaken by an important non-profit organization which works towards the social integration of people with disabilities in Spain. This organization has set out to evaluate its employment programs between the year 2000 and 2009.

This is a highly original initiative for the following reasons: Firstly, because the evaluation encompasses an extensive period (nearly 10 years) and aims to identify the effects of these interventions on the lives of people who have participated and on their family and professional environment. We have to bear in mind that the target population is a particularly sensitive group which requires regular and specialized supervision in order to achieve integration within the work environment. Only a qualitative approach can assume the changes experienced as a result of the programme in addition to its limitations and difficulties.

Secondly, because the evaluation has been established within the framework of a participative process of the stakeholders. This participation includes the people who benefit, their personal environment, business owners, the program’s technical personnel and the civil society most directly affected. It is outlined during a series of workshops which detail the design of the evaluation and enrich the interpretation of the results thus ensuring the usefulness of the recommendations for the organization itself.

As such, a participative methodology, known as 5x5 (five participants per five profiles in five territories) manages to include different visions in each of the phases of the evaluation. It also contributes to the institutional and motivational strengthening of the participants.

Keywords: Participatory evaluation methodology; European Social Fund; Fight against discrimination

Evaluation use by unintended users: The role of site-level stakeholders in large multi-site evaluations

F. Lawrenz
University of Minnesota, Educational Psychology, Minneapolis, USA

Although there has been extensive research on the use of evaluation, as well as about the involvement of stakeholders in local participatory evaluation, less is known about how stakeholders are involved in multi-site evaluations or how they make use of them. This project investigated the role of site level stakeholders in multi-site evaluations of four US National Science Foundation (NSF) programs. These were well-funded, large-scale, multi-site evaluation studies about the following programs: Local Systemic Change through Teacher Enhancement (mostly involving school districts), Advanced Technological Education (mostly involving community colleges), Collaboratives for Excellence in Teacher Preparation (mostly involving colleges), and Math Science Partnerships (mostly involving teams of colleges and school districts).

Over a four year period extensive case studies using data collected through artifacts, surveys of participants and interviews with evaluators and participants were developed for each of the four national evaluations and a cross case analysis was performed. The cross case analysis particularly examined the interaction of evaluation use and involvement in the evaluation by the stakeholders at the multiple sites. Because the evaluations were being conducted for the NSF, the site level stakeholders were not the directly intended users of the evaluation information. However, the case studies provide evidence of diverse evaluation use by the stakeholders and show that involvement in the multi-site evaluations was perceived differently than how involvement is reportedly perceived in local evaluations. The most important aspects of facilitating involvement and use appear to be careful and continuous communication, clear interfaces, awareness of life cycles at the individual, site, and program levels and opportunities for developing a community of practice. Descriptions of how to implement these ideas will be provided. These results extend existing theory about participatory evaluation to the context of multi-site evaluations and to use of evaluation by non-intended users.

Keywords: Evaluation use; Stakeholders; Multisite evaluation; Government program evaluation
Sectoral policy: Urban and regional development

European Evaluation of URBAN II programme—lessons for urban policy and lessons for future evaluators

T. Fox, J. M. De Vet
ECOTEC, Regeneration and Economic Development, Leeds, United Kingdom

This work presents the findings and lessons emerging the European Evaluation of the URBAN II programme (2006–2013). URBAN II was a 0.75 billion euro initiative which helped 70 cities across Europe deal with urban decline. The evaluation was undertaken by ECORYS on behalf of DG Regio at the European Commission in 2009/2010.

The evaluation methodology was a mixture of quantitative assessments of monitoring data, a broad assessment of all 70 programmes and then a detailed assessment of 15 case studies – the latter method being much more informative in terms of generating robust evaluation evidence. The evaluators found that many of the impacts coming from URBAN were either intangible (improved images of areas, higher levels of confidence among communities) or were simply not measured in any output, outcome or impact monitoring data. The case studies were therefore a significant part of the evaluation process – a key lesson for future evaluators to consider.

The results show that URBAN funding was used to support a wide range of activities – this was partly down to the flexibility and localized delivery methods that tended to put local stakeholders in charge of how funding should be used. This bottom up approach was quite different to normal Structural Fund approaches and this localized agenda was often seen as a critical success factor for programmes. The evaluation found the importance of ‘it’s not what you do its how you do it’ – with the critical success factors being around the URBAN ‘method’ (bottom up, integrated, partnership working) as opposed to the URBAN funding.

Keywords: Urban decline; Regeneration; European evaluation; Integrated regeneration

Future evaluation of JESSICA programme: a new millennium approach for structural funds

G. Pedini, H. Kramer
Deloitte Consulting, Public Sector, Neuilly-Sur-Seine, France

JESSICA (Joint European Support for Sustainable Investment in City Areas) is an initiative of the European Commission in cooperation with the European Investment Bank and the Council of Europe Development Bank, in order to promote sustainable investment, growth and jobs in Europe’s urban areas.

Considered as an innovative instrument, it is used for the first time in the 2007–2013 Structural Funds programming period. JESSICA offers the managing authorities of Structural Funds programmes the possibility to use interim payments to invest in Urban Development Funds (UDFs) through recyclable and recoverable financial mechanisms, essentially equity, guarantees and subordinated loans. Recovered funds may be reinvested through UDFs or returned to the managing authority to support other urban projects. Thus, JESSICA aims at providing the managing authorities a greater access to loan capital for the purpose of promoting urban development, including loans for social housing where appropriate.

Recently, detailed JESSICA feasibility and opportunity studies have been carried out in many Member States (Belgium, Czech Republic, France, Italy, Germany, Portugal, United Kingdom…) to help them determine how JESSICA could be best implemented in their respective context. In 2013–2014, the evaluation of JESSICA as a sector policy evaluation will be very challenging: can JESSICA be evaluated like any other structural fund programme (ERDF, ESF)?

Given the specificity of JESSICA and in a context where the programme has not been implemented yet, various key topics can be tackled and some questions anticipated:

– What are the differences and the common points between JESSICA and other structural fund programmes (objectives, involved structures, financial mechanisms/engineering…)?
– What lessons can be drawn out of the evaluation of traditional structural funds and how can they be useful for the future evaluation of JESSICA?
– Conversely, how can JESSICA help rethink evaluation of traditional structural funds (results-oriented culture, better use of resources…)?

Keywords: Evaluation; JESSICA programme; Structural funds; Urban development; Loan capital
The influence of investment subsidies on regional development
S. Krupnik
Jagiellonian University, Center for Evaluation and Public Policy Analysis, Kraków, Poland

Investment subsidies belong to the most popular means of public support for enterprises. Only in Poland, within next few years more than 2 billion euros will be spend to support small and medium sized enterprises. The support’s mechanism is unsophisticated. Companies have part of their investments paid from public money. Allegedly, subsidies serve the public interest by enhancing competitiveness of national economy. However, the empirical evidence do not support these claims.

The main research question of the presented project was following: what is the real influence of investment subsidies on decision making process of entrepreneurs?

The research draws on both program theory evaluation and argumentation analysis. In order to research the issue, grounded theory methodology was used. As a consequence, both methodological and data source triangulations were applied: both qualitative (eg. in-depth interviews) and quantitative (quasi experimental, content analysis) methods were deployed; viewpoints of manifold groups (i.e. entrepreneurs, decision-makers, journalists and experts) were reconstructed.

The study brought manifold findings. The main results of the investigation is that investment subsidies do not cause effects which they are very often thought to be. Within the study I followed the assumed logic of the intervention (chains of causes and effects) and found many flawed assumptions. They do not only indicate that the intervention is not worth implementation, but also explain why it is deployed.

*Keywords*: Regional development; Grounded theory; Investment subsidy; Program theory evaluation; Net effect
Evaluation and monitoring of interventions in transition economies

Impact Evaluation of Irrigation Projects in Transition Economy Countries

M. Mukhtarov
Project Implementation Unit, Baku, Azerbaijan

The present investigation being based on a standard evaluation approach goes beyond the indicators using modern economical and financial categories as structural components of the model.

The article describes the approach to impact evaluation within the project cycle in the case of World Bank project in Irrigational Sector of Azerbaijan on establishment of the WUAS (Water Uses Associations). The approach is time-invariant and can be applied in the long-term perspective after project completion.

The initial impact of the project has been measured through a baseline survey conducted in 2005 and the follow-up impact study in 2010. Surveys have been based on quasi-experimental and longitudinal design.

The baseline survey was conducted in those areas expected to benefit from project interventions, as well as some selected non-project control areas. The study included two phases, with the first one focused on collection of general data on population, farming types, cropping patterns, agricultural services availability, WUAs, agricultural machinery availability, and input supply units. The second phase of the survey focused on collection of more detailed information at farm level on land ownership, livestock and crop production, crop pattern, crop production cost, and irrigation practices. While the survey identified a set of quantitative indicators in detail, qualitative ones were developed to allow social impact monitoring of the project interventions, project outcomes, increased access to irrigation water and reliability in its delivery. The Citizen Report Cards approach was used.

Financial Analysis; crop and farm models were budgeted based on the surveys’ data.
Economic Analysis: based on the actual results being obtained by farmers in the project area according to the M&E Surveys, new evaluation scenarios were built for the Economic and Financial Analysis. The scenarios were built using FARMOD software. The 2005 surveys’ data helped to define the before project situation and the 2010 surveys data for the with project situation.

Bosnia Herzegovina Country developing and social inclusion strategy as bases for better quality of education and link with labour market

L. Maric
Directorat for economic planning BIH

Directorate for economic planning completed Bosnia Herzegovina Country development and Social inclusion strategies covering period through the end of 2014. In both strategies better quality of education and its link with labour market as well as employment of youth have been recognised as goals. As part of policy development and implementation processes Directorate wants to modernise social policy governance and develop new mechanism for improvement of evidence based policy making in Bosnia Herzegovina.

UNICEF is many years successfully involved in education activities in BiH and this year started the project related to education and employability. These processes are very much connected with Social inclusion strategy.

During previous cooperation DEP and UNICEF implemented Country Led Evaluation as joint activity. UNICEF is credible, helpful and flexible partner that understand the local needs and support the country to realise them as well as to implement global initiatives.

Now, DEP wants to establish new M&E mechanism for SIS with the focus on evaluation and wants to do it jointly with UNICEF. The first step would be new evaluation pilot process related to education and youth employability. The findings and recommendations of this Pilot exercise will be used as a model for other evaluation processes connected with social inclusion.

Keywords: Strategies; Partners; Cooperation; Development; Maric; Vukovic
A means for assessment of reasons behind progress in EU Structural Funds co-financed programmes: Simple Progress Overview Tool (SPOT)

B. Mezosi, J. Matolcsy
KPMG Advisory Ltd., EU and government services, Budapest, Hungary

SPOT is an easy-to-use tool for identifying, spotting potential reasons behind absorption or progress figures of Structural Funds related programmes (National Strategic Reference Frameworks, Operational Programmes).

SPOT is a tool used to identify the reasons underlying a figure describing, at a predefined point of time, the level of absorption (i.e. the value of reimbursed grants) of an intervention by breaking that figure down to constituent, meaningful factors (e.g. popularity, eligibility) that are likely to have a direct, substantial impact on the final result.

SPOT factors are individually meaningful numerical indicators generated and used in the course of the analysis the value of which is expressed in the form of ratios illustrating significant steps in the process of project selection. Multiplication of the SPOT factors for an intervention will yield the level of absorption of that intervention.

The final output of the SPOT analysis can also create a solid basis for estimating absorption perspectives of the programmes by building a simple mathematical model on the factors and defining possible scenarios (without interventions and with intervention to one or more than one factors).

The figure below summarises the ‘SPOT equation’, the basic concept of the tool.

As a result of the de-composition of absorption, the above-mentioned individual ratio values are calculated for the relevant programming level (i.e. Operational Programmes, measures, sub-measures). Some potential explanations to the low ratio values (constituting potential bottlenecks in the application system) spotted as a result of the development of the above SPOT equation are as follows:

Hypotheses developed upon the evidence based results of the SPOT equation can establish a sound basis for further investigation of reasons behind progress.

The current version of this tool is widely acknowledged and used in Hungary, Romania, Ireland, Malta and several other countries in the CEE region for progress evaluation purposes.

**Keywords:** Methodology; Progress; Absorption; De-composition
Joined up development evaluation

How can evaluation of development aid be made more coherent and joined up?

H. E. Lundgren, O. B. Tall, R. Picciotto, M. Segone, H. Wedgwood

Panel discussion, moderated by Hans Lundgren. Perspectives from leading actors in development agencies and from partner countries (4 persons, gender and nationality balanced).

OECD countries provide currently around $120 billion annually in development assistance. This is delivered through a variety of national development agencies, a great number of multilateral agencies, NGOs, and global and regional programmes. There is a clear demand for evaluation from parliaments, media and taxpayers in donor countries and from partners and citizens in developing countries who wish to see the results of the aid.

The supply response to the demand is a large number of evaluation reports produced by evaluation departments in development agencies. While joint evaluations have recently grown in number, individual evaluations are still clearly the overwhelming majority, and may only be part of an answer to the broader question posed by the session. In addition to centrally produced evaluations there is an unknown quantity of evaluations/reviews done by operational departments and in country.

An overview report will be made available by the OECD. Participants will draw on this and their own experience and perspectives in proposing ways and means how to achieve more joined up approaches in the various stages of the evaluation process, from planning to reporting on evaluation results. This also relates to issues of how to answer more aggregate questions on development effectiveness and aid efficiency (not working in silos) and how to achieve greater impact with an evaluation report on policy makers and on media.

Keywords: Evaluation of development
Evaluation cultures

Symposium: Evaluation Cultures

S. Speer¹, G. M. Hellstern (Chair)¹, K. H. Papazian¹, C. Jacot-Descombes¹, N. Stame²
¹Independent Evaluator, Wiesbaden, Germany
²University Kassel, Kassel, Germany

A Review of Frameworks for Evaluation: Emerging Trends within Sector-Specific Evaluation Cultures

Gerd-Michael Hellstern, University Kassel, Kassel, Germany
Sandra Speer, Independent Evaluator, Wiesbaden, Germany

The past decade has witnessed a significant growth of evaluation in many European countries at the European, National, Regional and Local level using different approaches and generating new ones. Since the first meticulous classic review by Stufflebeam in 1983 (revised 2000 and 2001), many more evaluation models have emerged, representing different schools of thought about evaluation. However, looking into practice some few models are dominating the landscape. Reflecting on current practice and the discourse on evaluation theory in the school sector, the labor market sector and urban development programs in Germany, we demonstrate how they are embedded in sector specific evaluation systems and cultures and how different approaches are addressing particular evaluation design challenges.

Riding the Celtic Tiger – forces shaping evaluation culture in Ireland in good times and bad

Boyle, Richard, Institute of Public Administration, Dublin, Ireland
O’Hara, Joe School of Education, Dublin City University, Ireland
McNamara, Gerry School of Education, Dublin City University, Ireland

The central theme of this paper is whether one can say there is such a thing as an evaluation culture or whether there are diverse cultures which change over time. Key issues in the creation of an evaluation culture are investigated, and in particular the impact of the rapid economic changes on the attitudes to and practice of evaluation are explored. The development of a culture of evaluation is examined in two areas: the civil service and education. While there are common aspects of an evaluation culture in evidence, in practice there are differing forces and values shaping the culture in these sectors. It is argued that it is in the interplay of the value conflicts between these forces that evaluation cultures are shaped.

Culture evaluation and multiculturalism

Katia Horber-Papazian, IDHEAP, Lausanne, Switzerland
Caroline Jacot-Descombes, IDHEAP, Lausanne, Switzerland

Swiss political culture is characterized by common values that allow the integration of a diversity of religious and linguistic cultures. In order to insure national cohesion, an important effort is asserted regarding the integration of national minorities, which implies the pursuit of consensus. It will be specified whether two characteristics of the Swiss federation – cultural diversity and democracy of consensus – affect the practices and the regional institutionalization of evaluation, the topics under scrutiny and the evaluator’s profile. Horber-Papazian/Jacot-Descombes also seek to specify if one can witness an increase of participative procedures allowing to integrating minorities.

The culture of evaluation in Italy: between distrust and responsibility

Nicoletta Stame:

The paper analyzes the tortuous ways by which an evaluation culture is developing in Italy. Contrasting favorable and less favorable aspects of institutions, practices and values that shape its political culture, the paper compares how it could be possible to overcome distrust and to promote an attitude of responsibility in three policy domains into which evaluation has been introduced: the University, the administrative reform, Structural Funds.

Riding the Celtic Tiger – forces shaping evaluation culture in Ireland in good times and bad

J. O’Hara¹, G. McNamara¹, R. Boyle²
¹Dublin City University, School of Education, Glasnevin Dublin, Ireland
²Institute of Public Administration, Senior Research Officer, Connaught Dublin, Ireland

A central theme in this paper is whether one can say there is such a thing as an evaluation culture in a country or organization or whether in reality there are diverse evaluation cultures. An interesting sub-text associated with this issue is whether evaluation culture(s) change over time. This paper explores key issues in the creation of an evaluation culture for example:

• what values get prioritised in practice?
• who are the international and national networks and groups who shape evaluation values and culture?
• the impact of economic rise and decline on attitudes to evaluation in the Celtic Tiger – specifically comparing and contrasting evaluation values during the Celtic Tiger era with those now dominant during the current financial decline.

The paper addresses these issues by examining the development of evaluation culture in two parts of the public service – the civil service and the education sector. In the case of the civil service the “value for money and policy review initiative” is explored and contrasting this in education the emerging “whole school evaluation system” is highlighted. This work builds on earlier research studying the emergence of evaluation culture in the public service in Ireland (McNamara, O’Hara, Boyle and Sullivan, 2009). It is argued that while there are common aspects of an evaluation culture, in practice there are differing forces shaping evaluation culture in Ireland, each with differing values at their core and that it is in the interplay of value conflicts between these forces that evaluation cultures are shaped.
Evaluation in context of persistent gender inequalities

Making evaluation matter in contexts of persistent gender inequities

K. Hay¹, R. Sudarshan¹, S. Zaveri³, R. Khanna⁴
¹IDRC, Evaluation Unit, Delhi, India
²Institute for Social Studies Trust, Director, Delhi, India
³Independent Evaluator, Founder member Community of Evaluators, Mumbai, India
⁴SAHAJ, Baroda, India

Women’s rights have been a focus of international development for almost half a century. Rights based frameworks’ and approaches have evolved and feminist analysis has produced new understandings of rights that are less abstract and more lived. Simultaneously, the field of evaluation has increasingly grappled with issues of context. This symposium argues that attention to context can bring focus to how rights (or inequities) are lived in particular settings. They symposium relates persistent gender inequities in South Asia to feminist and evaluation theory, and connects both to work on the ground. The paper by Katherine Hay argues that in the context of persistent gender inequities, evaluation must play a stronger role in distinguishing what works in shifting norms and inequities. Taking examples including women’s roles in decentralization and declining sex ratios, she argues that: gender programs and policies would benefit from deeper use of theory driven evaluation; many implicit theories underpinning such programs remain largely untested, and; attempts to test these theories are done in isolation. The paper by Ratna M. Sudarshan examines experiences of conducting evaluations of women centred, non-governmental development projects in India. The paper describes the context in which these evaluations have been carried out and examines the importance of capturing strategies of change to unpack both gendered theories of change and change itself. The paper by Sonal Zaveri discusses the need for evaluations to address contextual factors to understand the dynamic nature of settings. The paper argues that the increasing ‘commoditization’ of girls and women has implications on how and what we evaluate. The paper uses case studies on sex workers, children affected by HIV, and migrant women workers, to demonstrate how context shapes evaluation questions, approaches and use. The panel chair, Renu Khanna, looks across the papers to explore issues of power in context.

Keywords: Context; Gender; Theory based evaluation; Women’s rights; Feminist evaluation
Capacity in national context

The impact of the Structural Funds on development of evaluation capacity and culture in Poland

A. Januszkiewicz
Technical University of Lodz, Department of European Integration and International Marketing, Lodz, Poland

The European Union’s Structural Funds (SF) have an important impact on development of evaluation practice in European countries. In some Member States it was considered as a part of bureaucratic requirements and abandoned usually with termination of EU assistance. In others it was used as an opportunity to spread evaluation culture on different areas of public policy.

The main objective of this paper is to analyse and indicate factors influencing development of evaluation capacity and culture in Poland. The adopted framework of analysing the evaluation capacity is coherent with EVALSED approach. However, two kinds of factors are considered: external (EU evaluation model) and internal (national structure of SF management, legal, institutional aspects, administration culture, political attitude).

Evaluation practice in Poland has been strongly influenced by the realisation of the EU funds. It has been considerably developed in quantity and quality, first at national level (2004–2006), then at regional one (since 2007). For the demand side of evaluation capacity, the biggest progress can be assigned to setting formal institutional structures for evaluation and good quality of human resources engaged in evaluation at administration (knowledge, attitude and personal commitment). For the supply side, the biggest barrier is an insufficient number of evaluation experts. Evaluation practice outside the EU funds is still very limited. However, there is an intensive debate at national level to expand evaluation to overall public policy and administration practice. In Polish case external factors were very important to give a first stimulus to introduce evaluation practice, but internal factors are crucial to use this practice to build national evaluation capacity and culture.

Keywords: Structural Funds; Evaluation capacity

Challenges faced by French Evaluation Society to keep together a professional community despite unstable contexts

A. Fouquet
Société Française de l’Évaluation, présidente, 75013 Paris, France

How to keep on gathering the professional community of evaluators in a country where the way towards a culture of evaluation is going back and forth? The story of the French professional community is rather peculiar: it grew up upon different sources of legitimacy: the inter ministerial device for evaluation (created in 1990 and dismissed in 2002), then European structural funds (as elsewhere) and evaluation of specific policies contracted by the French State (national level) and Local Authorities following the Decentralization Process; more recently, (2008) the affirmation of Evaluation in the French Constitution and its subsequent reforms of Parliament and Cour des Comptes (i.e. somehow National Audit Office); and, new coming, last but not least, the proselytism of quantitative scholars (Random Controlled Trial via Paris School of Economics). To keep on federating the professional community in such a moving context, SFE has to lean upon new sources of legitimacy, such as the Circles of the Reform of the State (network of State civil servants), the building of specific networks of actors (elected decision makers, local civil servants…), sectorial groups, forums; all this has to be done without forgetting SFE’s own story and must in the mean time strengthen SFE’s values (debate on the Chartes).

Keywords: Professional community; National evaluation society; Ethics

Monitoring and Evaluation of Public Service Departments in South Africa

J. Mathe, T. Masangu, N. Gcwabe

1 Director: Programme Evaluation, Public Service Commission, South Africa – JabuM@opsc.gov.za
2 Deputy Director: Programme Evaluation, Public Service Commission, South Africa – ThokoM@opsc.gov.za
3 Deputy Director: Programme Evaluation, Public Service Commission, South Africa – NoqoboG@opsc.gov.za

This paper sets out to discuss the Transversal Public Service Monitoring and Evaluation System used by the Public Service Commission of South Africa to assess the performance of Public Service Departments in terms of compliance with the nine values and principles of good governance set out in section 195 of the Constitution of South Africa. The paper describes how the system works in terms of methodology and process and highlights key findings and recommendations made by the Public Service Commission during the past three years. It further assesses the impact of the system in terms of improving the quality of governance in the South African Public Service.

Keywords: Monitoring and Evaluation; Public Service; Compliance; Good governance
Evaluation and youth I

Evaluation in informal and non-formal learning settings

J. Galla 1, E. Gelabert 2, T. Scholze 3, C. Niesmann 4
1 die Berater, EU department, Vienna, Austria
2 BELIES Consulting, Director, Barcelona, Spain
3 Blinc Cooperative, Chairman, Göttingen, Germany
4 BELIES Consulting, Project department, Barcelona, Spain

Background
Informal and non-formal learning and the validation of competences gained in informal settings have attained popularity in the fields of youth, vocational training and adult education. As outlined by the EC's Directorate General for Education and Culture and the European Centre for the Development of Vocational Training, validating informal and non-formal learning is increasingly considered as a pathway to lifelong learning. While assessment focuses on the progress of the individual learner and is primarily a voluntary process, the interest in evaluating activities that lend themselves to informal learning on a project level has gained significant ground, too.

Objectives
The presentation will highlight the results and lessons learned of evaluating learning outcomes on individual and collective levels in three particular settings: (1) a volunteer cultural renewal project on community level, (2) volunteer practices in NGOs in the European Voluntary Service, and (3) an international community of practice of project managers. In these three case studies we applied methodologies and tools described in the evaluation system approaches LEVEL5 and Outcome Mapping.

Results
We use an action research framework to illustrate, appraise and reflect upon our experiences made in the application of the LEVEL5 and Outcome Mapping approaches in informal learning settings. Applying traditional assessment methodologies to learning are being criticized for being hardly useful in informal and non-formal learning settings that are characterized by dynamic and complex processes. The interests and motivations of the learners, their resources and social circumstances have to be taken into account throughout the assessment and evaluation process. A model for negotiating domains of competences and their meaning within the project context will be developed. A case will be made that results of evaluations of informal learning settings can be useful for both the learners and project managers as users of the evaluation.

Keywords: Informal learning; Communities of practice; Competences; Evaluation use; Validation

DG ENV's EYE – Incorporating the youth perspective into evaluation

G. Ellis, M. Pai
Gellis Communications, Strategic communications, 1050 Brussels, Belgium

Gellis Communications was contracted by DG Environment (DG ENV) to carry out an evaluation of its ‘Environment for Young Europeans’ (EYE) website. The EYE website was originally developed to raise youth awareness of environmental issues: although it aimed to be primarily ‘informative’ rather than ‘educational’, the key objective was to support Environmental Education.

The evaluation was carried out over a 9-month period, and was unique in that Gellis collected the opinions and usage of environmental youth websites from young people aged between 6 to 16 years old. The evaluation also included teachers, and assembled their ideas on how the European Commission could make better use of the Internet to communicate environmental issues to young audiences.

The evaluation used several research tools:

- Expert website audit
- Market analysis – Digital mapping of the competitive environment
- Best Practice Analysis – 7 cases
- 10 EU Member State surveys – online survey for 825 students aged between 11–16 years; 100 teachers; 94 one-on-one qualitative interviews with teachers; 109 Focus groups with children.
- Target Group Consultation – online survey of 501 teachers, youth workers and parents, and 591 students aged between 11–16 years old, 62 one-on-one interviews with teachers.

The new strategy was developed to align with the new communication challenges envisaged for the EYE website. It aimed to satisfy users, improve overall quality and content, and raise awareness of the website. Moreover, it aimed to assist DG ENV in identifying optimal target groups, structure, content, tools and features to meet the website’s predefined communication objectives; establish the best ways to achieve them; and ensure the website’s sustainability.

Gellis’ provided detailed strategic recommendations for the future development of the EYE website: The Commission intends to revamp the Europa website sometime in 2010 and will be taking Gellis’ analysis and evidence-based recommendations on board at that time with respect to Environment for Young Europeans.

Keywords: Strategic communications; Evaluation; Market research
Policy and Public Administration

A review of the linkages in the monitoring & evaluation of public administration: A proposal for an improvement model

M. Seabi
Office of the Public Service Commission, Governance Monitoring, Pretoria, South Africa

In South Africa the monitoring and evaluation of public administration practices in the Public Service is reaching a peak (climax) level in that national and provincial departments have now taken on board the institutionalisation of an M&E System. Despite the good legislative and policy frameworks in place to guide close monitoring and evaluation evidence confirms that the Public Service is far from achieving a satisfactory level of effective public administration (in accordance to Section 195–197 of the Constitution). The major challenge identified is the weak linkages of the M&E System processes in the Public Service with regard to information obtained for evaluation and the use of evaluation results. The interaction with public servants and/or evaluation users reveals that this was the major barrier to the contribution of M&E to effective public administration. The M&E System often lacks good and effective mechanisms for feedback and lessons sharing. The absence of mechanisms for strong linkages means that the M&E System does not do justice to improving public administration practices. A model on linkages is being proposed for a comprehensive M&E System, with clear process linkages. This model includes a recommendation for the establishment of a National Evaluation Policy, which will guide the linkages of the M&E System towards integrated efforts, elimination of duplication, quality evaluation and improved levels of feedback. The model further makes provision for the introduction of a national electronic Monitoring System that gives real time information for effective decision-making and programmes or projects implementation. This will be a System that provides for a web-based evaluation information system for sharing lessons across the Public Service.

Keywords: Comprehensive M&E System; Linkages; Real time information; Lessons; Effective public administration; Evaluation use

"Cross-programme Evaluation" as a new approach for evaluating programmes: The case of European Territorial Cooperation” programmes

R. Hummelbrunner
CEER Regionalberatung, Graz, Austria

Background
The main aim of cross-programme evaluations is to obtain learning effects through comparative insights on similar issues in a range of programmes. A comparative analysis of their experiences and challenges should lead to cost-effective and fast delivery of findings through economies of scale and time. Responding requests of several programmes, the INTERACT Point Vienna has commissioned two pilot projects for cross-programme evaluations of programmes for the Objective “European Territorial Cooperation” (ETC). The author is Team Leader of the evaluation on Operational Aspects, in which participate 14 ETC cross-border programmes in Central and South-Eastern Europe.

Objectives
The purpose of the paper is to describe the methodological approach as well as the experience gained with this cross-programme evaluation of Operational Aspects, which is currently carried out and should be finalized until July 2010.

Methods
The presented cross-border evaluation is needs-driven and guided by the interest of participating programmes. The topics were agreed by them and focus on three operational aspects along the programme cycle (Project Generation, Assessment and Selection). Most of the data collection is carried out by a team of national experts and since much of the information deals with qualitative aspects and requires tapping tacit knowledge of key actors, face-to-face interviews are predominantly used. Complementary to these interviews two on-line surveys are conducted to capture the perspectives of selected stakeholders (members of Monitoring Committees, supported projects). The validation of the analysis and the exchange of experience will take place in synthesis workshops with (three) sub-groups of programmes, by using the focus-groups method.

Conclusion/Application to practice
Cross-programme evaluations are a collaborative learning exercise and not a judgment of individual programmes. They cannot substitute, but should complement and inspire programme-specific evaluations – the insights derived from them could be used as inputs for programme-specific evaluations.

Keywords: Cross-programme evaluation; Comparative analysis; Learning-oriented evaluation; European Territorial Cooperation
Evaluation conflict and violence

A New Paradigm for Police in Rural Cambodia: The Adoption and Adaptation of “Community Policing”

I. Ramage, K. H. Ramage, E. Mazard
Domrei Research and Consulting, NA, Phnom Penh, Cambodia

What does non-violence mean, in practice, for national governments and international donors seeking to improve daily life in an impoverished, post-war setting? The implementation of an Australian supported “Community Policing” initiative in Cambodia provides some unique insights into these issues, in a cultural context that is more famous for its traditions of violence (in the 20th century) than its traditions of non-violence (perhaps to be found in the underpinnings of Theravada Buddhism).

An impact evaluation comprised of 1200 random household surveys, supported by a further 30 (non-random) in-depth interviews, examined both the real and perceived changes in rates of violence.

Although the initial purpose of the survey was to measure the effectiveness of the intervention itself, the findings reveal a range of important information as to how conflict resolution and ameliorating violence works (and doesn’t work) at the level of local village governance.

Specific findings deal with violence against women and the counter-posing issues of violent youth and the reactions of police and village authorities to them. These authorities are often in the position of managing conflicts (and forms of violence) that are not quite crimes; the eager adoption of the preventative paradigm of the “Community Policing” strategy reflects an awareness of the limitations of a purely punitive role.

The outcomes of this initiative (CCJAPIII) in several villages of Cambodia demonstrate the capacity of rural people to adopt and adapt new ideas (from a foreign source) and put them to good effect, with outcomes that can be measured by independent evaluation. This shows that a latent interest in reducing violence in the village, and even within marriages, can take on new impetus with a little political direction, in at least some cultural contexts.

Keywords: Crime; Policing; Violence; Post conflict

Evaluation of development programs in post-conflict situations: lessons learned and future challenges

K. Atanesyan, S. Markova
World Bank Group, Independent Evaluation Group, Washington DC, USA

International development community is assisting many developing countries and territories emerging from protracted internal and external conflict situations (e.g. Timor-Leste, Nepal, West Bank and Gaza) to tackle the challenges of state-building, reconstruction, social cohesion and human development. Post-conflict situations present specific challenges for designing and implementing a comprehensive development program: (i) volatile political situation, due to unresolved or “frozen” nature of many conflicts; (ii) lack of social cohesion necessary for nation building, resulting from internal conflicts (sectarian, ethnic, political, etc.); and (iii) extremely weak institutions and poor governance. Multilateral development institutions (UN, EC, World Bank) often find themselves in the position of leading the efforts of the development community. Presence of multiple donors, while positive from the perspective of availability of higher volumes of aid resources, can present an additional challenge from the perspective of aid coordination and effectiveness. Effective and credible evaluation of such programs can potentially produce important lessons to be learned for all involved parties: recipient governments, donors, and the civil societies. Another possible positive externality could be the entrenchment of the evaluation culture in nascent institutional environments.

This presentation will build on the experience of the past and on-going evaluations of country development programs, conducted by the World Bank's Independent Evaluation Group. It will look at several cases of joint donor efforts. One such recent country-case will be analyzed in greater detail.

The paper will identify: (a) a set of specific lessons learned; (b) a number of forward-looking recommendations, in regard to strategy design, implementation and evaluation methodology; and (c) a few specific challenges that will need to be addressed to design, implement and evaluate a comprehensive medium- to long-term development strategy in a post-conflict situation. Advice and guidance of conference participants are expected to be an invaluable input to the latter issue.

Keywords: Post-conflict; International development; Program evaluation; Institutional capacity; Aid effectiveness
Evaluating support to conflict prevention and peace building at global level
S. Wille1, E. Clerckx2
1European Commission, Joint Evaluation Unit External relations family, Brussels, Belgium
2ADE Consulting, Evaluation Department, Louvain-la-Neuve, Belgium

“The Joint Evaluation Unit of the External relations family in the Commission of the European Union is conducting an evaluation of EC support to conflict prevention and peace building activities in third countries. Two features call for a specific approach towards such an evaluation. First, few themes are as dependent on context as this one, so that designing an approach that allows drawing conclusions at a global level is particularly challenging. Moreover, the European Commission, as most of the other donors, aims at having an all encompassing, integrated approach when providing support to conflict prevention and peace building. This consists in a kind of paradigm shift where conflict prevention and development are closely intertwined. In this context, how can support to conflict prevention and peace building be evaluated without having the unrealistic ambition to evaluate development as such?

The presentation first outlines the frame set by the commissioning body, the constraints and inherent difficulties of this demand. The second part of the presentation explains the choices made, the approach proposed as well as the limits of this approach. The presentation will be made jointly by a representative from the commissioning body and a member of the evaluation team.

Keywords: Strategic evaluation; Conflict prevention; Complex evaluation; Peace building; Evaluation design

Assessing relevance of international development and humanitarian response at aggregate levels – example summary Child Wellbeing Outcomes from World Vision International

N. Morrow1, N. Mock2
1World Vision, Evaluation and Research, Viterbo, Italy
2Tulane University, Public Health, New Orleans, USA

Whether international aid contributes to sustainable development continues to be debated. Demands are growing for international Non-Governmental Organizations (NGOs) to demonstrate contribution to improved wellbeing outcomes in communities where they implement development and relief projects. A key assumption of many impact evaluations is that the programs were designed to be relevant to community needs. Yet demonstrating the relevance of a particular program with respect to community needs, NGO organizational policy, and organizational capacity is rarely documented. International NGOs often collect the essential information for understanding the relevance of their programming portfolio, but lack the tools or experience to analyze the question of relevance at aggregated levels. In 2009, World Vision International undertook a large thematic review of the relevance of their programming portfolio at the national, regional, and global level as part of an initiative to improve impact reporting. The paper describes the availability and quality of data that can be used to assess relevance. It presents several novel conceptual tools that were used to distill relevance assessment from participatory workshops and document review. The thematic review of relevance found mixed results and great scope to improve the relevance of programming through more effective use of routinely collected data. This study showed that impact and sustainability of programs might be increased through better design, emphasizing relevance to community needs and match between needs and NGO capacity.

Specifically, this paper contributes to the field of evaluation by:
– highlighting issues in assessing the key evaluation criteria of relevance,
– provides a multi-country perspective on collecting evidence, participatory analysis, and use of evaluation,
– explores issues of aggregation at different hierarchical levels of contemporary international non-governmental organizations,
– introduces novel participatory approaches to analyze relevance,
– discusses existing evidence for improved analysis of relevance,
– and, presents findings that imply a greater focus is required on the question of relevance.

Keywords: International Development; Relevance; Novel participatory approaches; International non-governmental organizations; Multi-country
Sectoral policy: Global Health (&Technology)

Evaluating Global Health Partnerships – A Shared Learning Approach

D. Brown
Performance Management and Evaluation, Paragon Review and Consulting Inc., Ottawa Ontario, Canada

Background:
Globalisation has exposed both challenges and opportunities for action. New organisations and networks such as the World Health Organization’s Global Noncommunicable Diseases Network (NCDnet) have been formed to address global health issues. NCDnet’s vision is to reduce risk, morbidity and mortality related to noncommunicable diseases and their shared modifiable risk factors through effective collaboration focused on achieving results in low- and middle-income countries.

The challenges of multi-sectoral collaboration are vast, even more so when they aspire to collaborative action. The role of evaluation is critical: new partnerships need to assess their progress, evaluate the results, and translate lessons learned into action.

Increasingly, the evaluation of global health partnerships/collaborations needs to be seen through the lens of learning from experience and shared learning.

Objectives
This symposium will:
1) provide an overview of the Global NCDnet, critical path, success factors and challenges;
2) demonstrate how evaluation processes promote learning;
3) describe NCDnet’s evaluation framework and indicators; and
4) discuss preliminary results.

Method
A strategic evaluation framework has been developed to assess performance of this multi-sectoral collaboration. The methodology uses a wholistic approach, incorporating a range of perspectives, methods and sources of information.

Conclusion
This symposium will draw out lessons that may be helpful to other organisations striving to create or maintain global partnerships.

If you want to go quickly, go alone; If you want to go far, go together. African proverb

Keywords: Policy; Strategic; Health; Global; Partnerships

Evaluating the contribution of Canada’s Global Health Research Initiative to global health goals and commitments

N. Diop, M. Cohen, C. Clemenhagen
Global Health Research Initiative (GHRI), International Development Research Center, Ottawa Ontario, Canada

Background
The Global Health Research Initiative (GHRI) is a partnership of five institutions of the Government of Canada: the Canadian International Development Agency, the Canadian Institutes of Health Research, Health Canada, the International Development Research Centre and the Public Health Agency of Canada. GHRI funds multi-disciplinary teams of researchers and research-users from developing countries and Canada to tackle global health priorities in low and middle-income countries (LMICs) such as maternal and child health, emerging infectious diseases, and health systems strengthening. GHRI-funded research addresses Canada’s nationally and internationally agreed goals and commitments in global health.

Objectives
GHRI evaluation allows assessing achievements, and informing Canadian agenda-setting and progress in priority areas for global health and global health research.

Methods
Evaluation relies on 1) a three-level ongoing evaluation. Levels 1 and 2 are addressed in GHRI evaluation strategy. They measure achievements, document processes, and examine the contribution of the partnership above and beyond what each partner can do alone. Level 3 consists of funded projects’ evaluations; 2) an external formal review focusing on GHRI strengths, weaknesses and partners’ perception of value-added.

Indicators are elaborated from a conceptual framework. Concurrently, qualitative methods used include review of literature, repositories and web-based resources; focus group discussions and in-depth interviews with key stakeholders; and exchange with other donors in global health. Case studies and special studies complete the evaluation strategy. The formal review is based on interviews with key informants.

Results
Selected results of the ongoing evaluation will be presented, highlighting contribution to global health in LMICs. Key findings of the formal review will also be presented, highlighting contribution to Canada’s goals and commitments.

Conclusion
GHRI is a unique partnership. It is anticipated that evaluation findings will nurture collaboration between partners, inform Canadian approaches to global health research and stimulate investment in global health research.

Keywords: Global health research; Global health; Canada; Low and middle-income countries; Partnership
Evaluating the impact/contribution of ‘improvement’ bodies in the health sector: a pragmatic programme theory approach

E. Hamilton
The Tavistock Institute, London, United Kingdom

In the UK, there are a number of arms-length organisations working to improve the quality of healthcare in the NHS. These have no direct authority over NHS professionals or bodies. Their work therefore involves collaboration to innovate and influencing to bring about spread/ adoption of these innovations, in order to achieve improvements in quality – the intended impact.

For these organisations, it is important for them to gain an understanding of the success (or impact) of their work: Can the innovations bring about the intended effects? Are the innovations being adopted? To what extent are their intended outputs/outcomes occurring?

This presentation will investigate various difficulties, such as: clarifying the programme’s intended outputs/outcomes in the context of an evolving solution/evaluand; setting output and outcome measures; utilising the concept of contribution rather than impact, and thereby recognising the limited scope of possible impact (and evidence of it) in a crowded policy sphere.

In this context, the option to construct a programme theory will be investigated, along with the various benefits and challenges. On the basis of this, various intended outputs and outcomes can be clarified, and measures assigned. The evidence required to measure these will be discussed. While this approach may not produce evidence high on the Maryland Scale, it can be a useful and meaningful start to constructing a useful programme, understanding the intended effects of activity, be an aid to improved effectiveness as well as providing a basis to understand an organisation’s contribution to longer-term outcomes.
International development

Methodological issues linked to the evaluation of development cooperation policies: the case of decentralized aid policy in Spain
E. Malagon¹, K. Unceta², I. Labaien³
¹HEGOA – University of the Basque Country, Applied Economics V, San Sebastian, Spain
²HEGOA – University of the Basque Country, Applied Economics I, San Sebastian, Spain
³HEGOA – University of the Basque Country, Research, San Sebastian, Spain

This paper studies some of the problems that are related to the evaluation of aid development policy, specifically in decentralized context, and that affect the feasibility of the evaluation of these policies. During last decade, regions and other lower government levels have become active agents of aid development policies, defining and designing their own agenda. Once the relevant differences among project evaluation and policy evaluation are highlighted, our work analyzes in the difficulties found in a research focused on the aid development policies of the Spanish regions (Comunidades Autónomas). Finally, the paper concludes with some proposals for the design of policies that could improve their evaluability.

Keywords: Aid policy; Spain; Decentralization

Why will we never learn? Different tactics for fighting against the crime of learning in International Cooperation. The Spanish case
C. Rodríguez-Arizá¹, C. Cordero¹
¹FIIAPP/Evaluation Unit of the Directorate General of Development Policy Planning, Dgpolde, Madrid, Spain

The Final Report of the Evaluation Gap Working Group was “When Will We Ever Learn? Improving Lives through Impact Evaluation”. As a means to provoke discussion, we changed the question: “Why will we never learn in Spanish Cooperation?”

The aid effectiveness Agenda is promoting mutual accountability and managing for development results. This process makes it more urgent improving the information management in development organizations. In this context, the quality of information arising from monitoring and evaluation of interventions implemented by development organizations depends largely on how the principles of Participation, Learning, Transparency and Utility are understood, defined, implemented between and within each organization. The understanding, definition and implementation of these principles have to take into consideration (1) the structural constraints of the “aid chain” but also (2) organizational aspects such as institutional and legal framework, motivation and organizational capacity. However, the inertia and the resistance to change are evident. As Britton (2005) says, many development organizations seem to operate as if they were implementing strategies of prevention of the “crime” of organizational learning. So these organisations do not seem to provide neither the reason (Why?) nor the means (How?) nor the opportunity (When?) for such a crime. Especially in the context of the Spanish public administrations, the complexity of reversing this vicious cycle (to another virtuous) implies an integrated approach within and between development organizations on several fronts, both structural and organizational.

This document is the result of a Ph.D research on information management in International Cooperation and the publication “The case of the Departments of the Spanish Public Administration working through NGDOs. A challenge and a commitment associated to the Effectiveness Aid Agenda”. The central aim of this article is to determine what information, why and how to utilize the information from monitoring and evaluation in Public Departments.

Keywords: Learning and Transparency; Paris Declaration; Managing for Development Results and Mutual Accountability; Participation and Power; Effectiveness Agenda; Knowledge and Information Management

R. S. Parker
IEGSE

Almost a third of all World Bank projects deal with water. IEG recently completed an evaluation of World Bank lending for water. The report (released on March 22, 2010) examines inter alia hydropower, floods and drought, water supply and sanitation, watershed management, rivers, lakes, coastal zones, inland waterways lending between 1997 and 2007, analyzing the full universe of 1864 relevant projects. Together, these projects represent Bank financing of US $54.3 billion. The findings informed a revision of the World Bank water sector strategy. Methods included country and regional case studies, preparation of issue papers on specific water topics, and the creation of a master database which incorporated data from project files, case studies, background papers, and individual project level evaluations.

IEG finds that project performance has been improving steadily against stated objectives. Water has also been successfully integrated with many other sectors. The Bank has balanced investments in infrastructure with improvements in the institutions that manage and allocate water. However, the environment’s needs for water still get short shrift. Countries rarely want to borrow to rescue an aquifer, or to restore protective marshes and wetlands, for example. IEG finds that given their economic and strategic importance, groundwater conservation, coastal zone management, sanitation, and environmental restoration have been underemphasized, and this is at least partly because data on their economic importance is poor and unlikely to improve without better monitoring.

The IEG report points out that the Bank needs to tackle several under-emphasized issues given that borrowing countries will face heightened challenges in the coming decades. Faced with mounting shortages, the international community must find new ways to support countries in managing their water resources. The challenge is to meet today’s water needs addressing growing water stress – manifested in shortages projected in the order of 40 percent by 2030.

Keywords: Sector evaluation; World Bank; International development; Infrastructure; Environment
Evaluation of anti-tobacco policies

Mass media campaigns for public education and to generate support for tobacco control policies in high-prevalence countries in Europe

R. Perl¹, N. Murukutla¹, I. Morozova¹, S. Mullin¹
¹World Lung Foundation, Communications, New York, USA

Evaluation of anti-tobacco policies

Mass media campaigns for public education and to generate support for tobacco control policies in high-prevalence countries in Europe

R. Perl¹, N. Murukutla¹, I. Morozova¹, S. Mullin¹
¹World Lung Foundation, Communications, New York, USA

WLF symposium session 1

Tobacco is the single most preventable cause of death across the world and its impact across Europe is no exception. When it comes to tobacco control, the stakes are high in Europe because much of the world looks to Europe as a leader in public health policy. Public education through mass media campaigns are a proven strategy that can reduce tobacco use, support smoke-free policies and change social norms and behaviors around smoking.

Over the past three years, the World Lung Foundation (WLF) has been supporting governments in thirteen countries, including four in Europe that have been identified as “high priority” based on smoking prevalence and income, in implementing scientifically rigorous mass media campaigns to promote tobacco control initiatives and encourage smoking cessation.

In this introductory session we provide an overview of the approach that WLF has adopted to measure campaign impact. Specifically, we describe the “social marketing” approach to the development and evaluation of the tobacco control mass media campaigns; the evaluation study designs and key indicators used to measure campaign efficacy and impact; and, some of the broad trends observed through the application of these principles in our campaign evaluations.

Subsequent sessions will describe recent WLF assisted campaigns and their evaluations in Russia, Ukraine, Poland and Turkey.

Keywords: Social Marketing Campaigns; Tobacco Control Campaigns; Public Health Campaigns; Russia; Ukraine

Evaluating an iconic social advertisement in Russia’s largest city

T. Voylokova¹, N. Murukutla¹, I. Morozova¹, R. Perl¹, L. Stebenkova¹, V. Kochetova¹, J. Baskakova¹, A. Kotov¹, S. Mullin²
¹WCOM, Research, Moscow, Russia
²World Lung Foundation, Communications, New York, USA

WLF symposium session 2

Background

The share of smokers in Moscow is among the highest in the world (60 % of men, 40 % of women) making tobacco control a critical health policy priority as it plays a significant role in Russia’s population decline.

In November – December 2009, the Moscow Duma, assisted by WLF, set a new standard for tobacco control advertising in Russia by adapting and airing an iconic advertisement that graphically depicts the serious health harms of smoking. The advertisement ran for eight weeks on television, billboard, and in newspapers. The primary objective was to increase awareness and change attitudes towards smoking.

Methods

The campaign was evaluated with a post-campaign in-person household survey of 1000 smokers and recent quitters in Moscow between the ages of 18–45 years. Campaign impact was assessed through logistics regressions that compared those exposed vs. not exposed to the campaign.

Results

65 % of respondents recalled advertising about the harms of smoking in the two months of the campaign. 50 % of respondents specifically recognized the campaign ad and most often recalled seeing the campaign ad through TV (63 %). 57 % agree that it was personally relevant. 47 % of smokers said it made them feel concerned about smoking. “Exposed” vs. “nonexposed” smokers were more likely to know smoking causes serious medical illnesses (74 % vs. 57 %, OR = 1.6), respiratory diseases (83 % vs. 68 %, OR = 1.7), heart attacks (56 % vs. 38 %, OR = 1.7). Exposure to the campaign also generated significant public discussion about the health consequences of smoking, an important precursor to behavior change.

Conclusions

This evaluation, demonstrated that hard-hitting public health education campaigns can be impactful in Russia and has paved the way for sustained programming critical for ultimately affecting prevalence rates.

Keywords: Tobacco Control; Mass Media Campaigns; Social Marketing Campaigns; Russia; Public Health
Using mass media communications to create support for smoke-free law in Turkey
N. Tunali, S. Hamill, N. Murukutla, T. Erdebir, S. Mullin
1 Strategic Focus Research Training and Consultancy, Research, Istanbul, Turkey
2 World Lung Foundation, Communications, New York, USA
3 Yesilay, Tobacco Control Project, Istanbul, Turkey

WLF symposium session 5

Background
In July 2009, Turkey implemented its landmark 100% smoke-free legislation, extending smoke-free environments in bars, restaurants, cafes and teahouses. From July to September 2009, WLF assisted its grantee, Yesilay, in implementing a “Smoke-Free Istanbul” campaign on TV and other media designed to educate the public on the harms of second-hand smoke (SHS) exposure and garner support for the law.

Method
Campaign efficacy was evaluated through two surveys: a pre-campaign (and pre-law) survey that was conducted in June and a post-campaign survey that was conducted in October 2009. Both surveys used the in-person household survey method to create representative samples of adults in Istanbul aged 18 – 50 years (Ns = 1,006 and 1,009 respectively).

Results
Support for the smoke-free legislation was already very high during the pre-campaign period with over 82% of Istanbulites believing that the ordinances were justified because of the benefits they would bring to the public. This general support for the law increased from the pre-campaign to the post-campaign period with greater numbers believing that the ordinances would help smokers quit (50% vs. 55%), fewer perceiving the ordinances as unfair to smokers (19% vs. 15%), and greater numbers believing that the law would be applied fairly (42% vs. 60%). There was also significant increase from pre- to post-campaign period in people’s knowledge of the association between SHS exposure and lung cancer (94% vs. 99%), heart disease (93% vs. 99%) and lung diseases in children (96% vs. 99%). Attitudes towards smoking also shifted in this period with greater numbers in the post-campaign period believing that Turkish society disapproves of smoking (56% vs. 63%). The evaluation indicates that the campaign increased both knowledge of tobacco’s harms and support for smoke-free laws.

Keywords: Turkey; Tobacco Control; Mass Media Campaigns; Social Marketing; Public Health Campaigns

Using mass media campaigns to educate and gain support for governmental tobacco control initiatives in Ukraine.
D. Krakovich, I. Morozova, R. Perl, N. Toropova, H. Hopko, N. Murukutla
1 Kiev International Institute of Sociology (KISS), Consultancy, Kiev, Ukraine
2 World Lung Foundation, Communications, Moscow, Russia
3 World Lung Foundation, Communications, New York, USA
4 Advocacy Life Center, Policy, Kiev, Ukraine

WLF symposium session 3

Background
Smoking prevalence among Ukrainian men is among the highest in the world (67% men, 20% women), and is increasing especially among less educated, and lower income men. Among women it’s the more affluent, highly educated women that have the highest smoking rates. And despite laws, 53% of the population has been exposed to secondhand smoke (SHS).

From December 2008 to May 2009, World Lung Foundation and partners ran a campaign using two complementary advertisements that graphically depicted the systemic health harms caused by smoking and secondhand smoke exposure. The primary objective of this national campaign was to increase awareness of smoking and SHS harms, shift public attitudes, and gain support for government tobacco control initiatives.

Method
Campaign impact was measured through a nationally representative post-campaign survey of 904 randomly selected individuals aged 18–50 years. by KISS. News Media was also monitored.

Results
The campaign ad’s were seen on TV by 32% of Ukrainians. The vast majority who saw the ad’s rated them as understandable, believable and discomforting. The ad’s prompted nearly 1 in 2 smokers to consider quitting. Those who were exposed to the ads vs. unexposed had better knowledge of the impact of smoking on serious illnesses, including heart disease, reproductive issues, and stroke. The campaign also generated over 80 news stories on various media, including a segment on a popular talk show.

Conclusion
This campaign generated a great deal of media attention. The messages were perceived as credible, persuasive and appear to have increased knowledge. In the months subsequent to the campaign the Ukrainian government passed legislation to increase taxes and added graphic warning labels to cigarette packs. It is plausible that the campaign, along with other critical tobacco control efforts, helped build momentum to support the government’s initiative.

Keywords: Ukraine; Mass Media Campaigns; Public Education Campaigns; Public Health Campaigns; Social Marketing Campaigns
Cigarettes eat you alive: The impact of a mass media campaign on knowledge, attitudes and behaviors of adults in Poland

A. Fronczak 1, K. Przewozniak 1, K. Suszek 1, E. Lata 1, N. Murukutla 1, R. Perl 1, I. Morozova 6, A. Koziel 7, M. Cedzynska 8, W. Zatonski 9

1 Ministry of Health, Vice Minister, Warsaw, Poland
2 Health Promotion Foundation, Scientific Research, Warsaw, Poland
3 Ministry of Health, Press and Promotions, Warsaw, Poland
4 Chief Sanitary Inspectorate, Public Health and Health Promotion, Warsaw, Poland
5 World Lung Foundation, Communications, New York, USA
6 World Lung Foundation, Communications, Moscow, Russia
7 World Health Organization, Country Office, Warsaw, Poland
8 Cancer Center and Institute, Cancer Epidemiology and Prevention, Warsaw, Poland
9 Health Promotion Foundation, President, Warsaw, Poland

WLF Symposium session 4

Each year smoking kills almost 70,000 Poles and second-hand smoke (SHS) exposure causes over 1,800 deaths among non-smokers. In 2009, the World Lung Foundation assisted the Polish Ministry of Health, other Polish public health institutions and the Polish Media Campaign Collaborative Team, to execute a national education campaign aimed at increasing awareness of tobacco smoke hazards, in an effort to change public attitudes towards smoking and gaining public and media support for up-coming stronger smoke-free legislation. From November to December, two complementary graphic advertisements ran on all major Polish TV stations. A national quitline supported smokers who were encouraged by the campaign to quit.

Campaign efficacy was evaluated through nationally representative baseline and post-campaign face-to-face questionnaire surveys of adults conducted in July and December 2009 respectively (Ns = 1000 each). The surveys measured the proportion of the population exposed to the campaign, their reactions to the campaign messages, and changes in knowledge, attitudes and behaviors. Volume of calls to the quitline were also used to evaluate the campaign’s impact.

According to the survey data, 83 % of Poles saw the campaign ads at least once and 64 % at least three times; which is considered an ideal amount. Half of the smokers exposed to the ad(s) said that it made them more concerned about the effects of their smoking on their family’s health, more likely to stop exposing others to their cigarette smoke, and more likely to quit smoking.

Since baseline, there was an increase in knowledge about the harms of smoking and SHS exposure, including its impact on asthma in children and stroke. And there was a concomitant increase in calls to the quit line during the campaign period when compared to the prior months of 2009 and when compared to the same period in 2008.

Keywords: Poland; Mass Media Campaigns; Tobacco Control; Social Marketing Campaigns
Monitoring and evaluating a complex global campaign

Oxfam GB’s experience monitoring and evaluating a complex global campaign

S. Starling, E. Otero, C. Hutchings, B. Cugelman

1 Oxfam Great Britain, Campaigns & Policy Division, Oxford, United Kingdom
2 Leitmotive, Consultoria Social, Seville, Spain
3 Oxfam Great Britain, Programme Performance & Accountability, Oxford, United Kingdom
4 AlterSpark, Toronto, Canada

Oxfam GB is an international NGO, working to overcome poverty and suffering through development, humanitarian and campaign programmes in 63 countries globally. Over the past few years, Oxfam GB has been working to improve how it monitors, evaluates and learns from its programmes, including its campaign and advocacy work. As part of this process, we have piloted a number of different approaches to M&E in the Climate Change Campaign – a large, high profile, complex campaign, working with national, regional and international partners to influence policy and public opinion in the UK, and more than 10 priority countries in the global south, in support of a fair, safe and binding global climate change deal.

Drawing on the literature and lessons around ‘Developmental Evaluation’, as well as our own experience, we have piloted ways of gathering high-quality evaluative information that supports real-time learning, and can inform decision-making processes during the campaign. This focus on action-oriented learning was complemented by an independent evaluation, commissioned to provide a rigorous, credible assessment of the results of the campaign.

Oxfam GB’s experience to date suggests that while some ‘traditional’ elements of M&E are helpful in advocacy work, a greater focus on light monitoring systems able to support real time learning is necessary. A level of accountability comes from the independent external evaluation. Crucially, our experience demonstrates that advocacy evaluation of even large and complex global campaigns is a) possible, and b) valuable.

The purpose of this session will be to share and discuss lessons emerging from our experience on the institutional and methodological challenges of assessing complex global campaigns. The session will be jointly presented by Oxfam GB M&E advisors and representatives of the external evaluation team, Leitmotive and AlterSpark. Efforts will be made to engage representatives from the South Africa and Bangladesh national campaigns as well (covered as case studies in the evaluation).

Keywords: Campaign; Advocacy; Climate Change; Developmental Evaluation
Contribution to policy and good governance

Performance auditing: contributing to accountability in democratic government

T. Ling¹, J. Lonsdale²
¹RAND Europe, Evaluation and Audit, Cambridge, United Kingdom
²National Audit Office, VFM development, London, United Kingdom

Along with Peter Wilkins, Lonsdale and Ling are editing a book to be published by Edward Elgar later this year on this topic. In this presentation, Lonsdale and Ling will draw upon the content of this book and outline the ways in which it can be said that Performance Auditing contributes to democratic accountability, how these hopes match with the reality of what public audit bodies actually do, and how this may be changing in the current circumstances. Drawing on concrete examples from performance auditing in a variety of settings, the potential contribution and actual performance will be assessed and illustrated. Examples will be taken from across countries with developed audit systems and with a particular focus on the experiences of the Netherlands, UK and US. The presentation will conclude with some considerations on future developments. Both Lonsdale and Ling have both worked closely and in auditing bodies and have published widely in this area and this presentation will build on and develop their position.

Keywords: Evaluation; Performance audit; Democracy; Accountability

Local audit committees in The Netherlands

D. Hanemaayer
B&A Group, Cbe, Oegstgeest, Netherlands

Local audit committees in The Netherlands

By law municipalities in The Netherlands have all a local audit committee. These committees support the chosen town councils in implementing there controlling tasks. Mostly these committees have about four to six members, partly members of the town council, partly experts (mostly in public finance). The members spend two to four days a month for committee-work, and there is a professional staff of one to three people. Research project are mostly commissioned to private research institutions.

On the one hand the local audit committees are a very welcome strengthening of the local evaluation capacity. On the other hand, almost all the committees have chosen a kind of formalistic, process-oriented evaluation which is hardly relevant for local policy.

Starting point in most evaluations of local committees is the policy cycle: in short policy goals – implementation of policy – realization of goals. The interpretation of this cycle by the committees is limited and formalistic. Research question are for example: are policy goals formulated, are these goals smart, is the town council informed about these goals, is there output information, is the town council informed about the output, and so on. So all questions about the technical quality of the policy process.

What is absolutely missing in the committee-work are important questions about the social impact of local policy. Which means questions which can contribute to enhancement of impact of local policy. We appeal for a drastic change in the research focus of the local audit committees.

Keywords: Audit committee; Formalistic evaluation
Evaluation and youth II

Navigating uncharted waters: Capturing the voices of children and youth in evaluation

D. Belanger
Proactive Information Services, Associate, Winnipeg, Canada

Background and Objectives
At first glance, it may appear evaluations of interventions to support gang-involved youth in Canada and young Roma children in East Central Europe would have nothing in common. However, these evaluations have to grapple with many of the same questions.

- How do the evaluators respectfully collect the voices of children and youth?
- How do they overcome trust issues inherent in working with vulnerable populations?
- As evaluations involving youth people move from a needs-based to a rights-based environment, what might evaluators have to consider or add to their toolkit in order to ensure the inclusion of those traditionally disenfranchised?

Methods and Results
Using illustrations from a number of evaluations in Canada and Europe, this presentation will explore data collection methods with children, youth, and those with emerging literacy in national official languages. The presentation will also discuss ethical considerations when eliciting the voices of those less frequently included in the evaluation process.

Conclusion
From a social justice perspective, the inclusion of those traditionally without a voice emphasizes the role of evaluation in creating a more just environment for all.

Keywords: Children and youth; Visual data collection; Social justice perspective

Designs and methodological challenges in evaluation of services for marginalized youth

K. J. Johansen, G. Hyrve
South Trøndelag University College, Ahs, Trondheim, Norway

Lately we have been working with evaluations aimed towards marginalized youth. Some of them have been addicted to drugs, some come from different child behavior institutions, some have been involved in crime and some have had psychiatric problems i.e.

In this presentation we will discuss our experiences with different evaluation designs and methodological challenges aimed towards services were the consumers are youth. We will sum up our findings in different projects we have carried out when it comes to collection of data and get into dialogs with different parts involved.

We will present concrete examples from evaluation projects were we will bring forward design, processes and results. Especially we will focus on an evaluation of an Am-car center used both as prevention of more problems and as rehabilitation of marginalized youth. This project is cooperation between the local municipality including school and child care, parents, police and the Am-car club. The evaluation has shown interesting results about how important new and valued roles are in helping youth in their rehabilitation process. We are in a dialog with central authorities on evaluation of new similar projects.

We will discuss different approaches connected with methodological questions. What is the strength of using different approaches towards marginalized youth? Why and how is it important to involve the youth in the evaluations processes? How can we use information from evaluations in constructing a new practice?

Keywords: Evaluation design; Evaluation of services; Rehabilitation process; Methodological challenges; Marginalized youth
Bollenti Spiriti Urban Laboratories: new life’s chance for young?

D. Morciano, A. F. Scardigno
University of Bari, Psychology (Sociology section), Bari, Italy

Background
In October 2009 the Puglia Region in the South of Italy committed to the University of Bari (Department of Psychology, Section of Sociology) an evaluation of “Urban Laboratories”. This programme is a part of a wide youth policy called “Bollenti Spiriti” and has been recognized as a good practice at European Year of Creativity and Innovation in 2009. It tries to face the crisis of traditional educational agencies (e.g. family, school, political parties) by promoting the participation of youth. Many priorities of EU youth policy are taken into account like enhancing youth participation, improving informal education and promoting youth work. Started in 2006, the programme has funded the refurbishment of 150 abandoned buildings. A number of 74 projects has been designed in a wide range of thematic activity (e.g. arts, performance, training, young entrepreneurship, local tradition, new technologies).

Objectives and methods
This ongoing evaluation has been oriented both by the theory-based (Weiss C., 1997) and the realistic approach (Pawson R., Tilley N., 1997). It was made from October 2009 to March 2010 by focusing on some case studies of Urban Laboratories. Evaluation research coincided with participation of youth in the selection of managing partnerships. Then evaluation aims at improving the participation process of youth towards the objectives to create new educational environment based on active and cooperative learning.

Results
Evaluation has identified some mechanisms related with context which are leading or not leading to create new informal learning environment. This mechanisms could generate new life’s chances for youth (Dahrendorf R., 2003) and break down the deterministic influence of its disadvantage situation. For example, urban laboratories are creating new forms of social capital by which youth can access to new opportunities of social and work inclusion (e.g. peer relations, outside network, relations within profit, no profit and public sector).

Keywords: Inequalities; Case studies; Innovative policy; Youth; Participation
Participants and co-production IV

Participatory evaluation of mental health day centres in the city of Rome: a project analysis

M. Spera 1, A. Rissotto 1, E. Cola 2
1 CNR – National Research Council, Institute of Cognitive Sciences and Technologies, Roma, Italy
2 National Research Council, Institute of Cognitive Sciences and Technologies, Roma, Italy

The contribution aims to analyse the project “Participatory Evaluation of Mental Health Day Centres in Rome” using as a key to interpretation the criteria proposed by Guba and Lincoln (1989) within the framework of Fourth Generation Evaluation.

The project was implemented by Institute of Cognitive Sciences and Technologies (National Research Council) and commissioned by the Department of Social Policies of Rome. It began in 2004 and was concluded in March 2010. The aims of the project were: to build up skills and tools to evaluate the work done, promote the role of the day Centres within the service network and the district, and support the motivation and project skills of the health and social workers.

The project has involved 26 semi-residential social health services. Several stakeholders participated, including: 40 managers, 95 practitioners, 80 users, and 60 family members. It entailed the implementation of several different evaluation processes regarding: quality of the services, organization, representations of the service, the planning of activities, training, communication and the sharing of results and products.

The project analysis will be based on the ‘authenticity criteria’ proposed by Guba and Lincoln: fairness cuts across the entire project; ontological authenticity and educative authenticity refer in particular to the explicit statement and comparison of the stakeholders’ views on quality and the mission; catalytic authenticity may be traced through the changes in the Centres’ planning modality and in the relationship with the other services and the district; tactical authenticity refers to the empowerment of the Department (commissioning agency) and the Centre practitioners.

The proposed analysis will highlight benefits and risks involved in a participatory evaluation project, and will discuss also the complexity of the evaluative design, which is built up through a comparison among the different stakeholders, and the importance of the temporal dimension.

Keywords: Participation; Stakeholders; Fourth generation evaluation; Mental health; Day centres

An evaluative approach from actors perspective: DOTS experience in Rio de Janeiro

M. Cruz 1, G. Cardoso 1, P. P. Chrispim 1, D. Abreu 1, P. Decotelli 1, J. Borenstein 1, E. Moreira dos Santos 1
1 Fundação Oswaldo Cruz/ENSP, LASER/ENSP, Rio de Janeiro, Brazil
2 Universidade Federal do Rio de Janeiro, Hospital Universitário Clementino Fraga Filho, Rio de Janeiro, Brazil

Objective:
The purpose of this study was to evaluate DOTS (Directly Observed Treatment Short-Course) strategy intervention for tuberculosis in an area of the city of Rio de Janeiro with high prevalence of the disease. Two Health Units, with different types of health system structure (one with health family strategy implemented and the other one with a traditional model) were selected in order to compare acceptability, accommodation and compliance to guidelines, presuming them as being associated to adherence.

Methods:
The study design was case-study, with a mixed-method approach in which phenomena are investigated in real life context by using direct observation, semi-structured interviews and documents analysis. Social representation theory was used in order to understand the meaning patients, their families and health professionals give to DOTS treatment, so we could describe and analyze which factors contribute and which are obstacles to patients adherence, considering health service organization and health professional-patients relationship dynamics.

Results:
In 63 interviews, field observation and document analysis conducted we identified different forms of access to DOTS. Both health professionals and patients, need to follow guidelines of treatment, which can be viewed as a limiting factor or facilitator, according to the representations about the disease and it’s treatment, in addition to situations presented: poverty, drug use and low family/social support. Other difficulties are related to protocol’s follow up itself and the irregularity in bus tickets offer.

Conclusions:
Factors such as reception, trust, respect and convenience, between patients and health professionals, as well as geographic accessibility, appear to be key to adherence, regardless the type of health unit structure.
Facilitating individual goal setting in group-based intervention

O. S. Kristensen

Although much research has been done on effective groups or institutional interventions, we still have much to learn about the constitution and facilitation of effective groups or institutions. During the past decade major differences in outcomes between different kinds of institutions have become apparent, and the specific practice that creates a particular outcome needs to be addressed. One of the major issues is the conceptualization of how different elements of a group affect persons in groups. This relationship between the group and the individual has been studied in terms of components, configurations or action nets, and these concepts may lead us to a new insight into the motivational aspect of an intervention. In this paper I will elaborate on these conceptualizations based on a study of a group intervention project for unemployed persons. Over the past 15 years, almost 600 persons have participated in the project. Today the project offers various workshops as well as social and psychological counselling and includes 31 participants. The participants experience severe social, psychological or psychiatric problems. A number of interviews with the participants are conducted, based on foto-eliciting techniques and a questionnaire about the participant’s current concern. These interviews were supplemented by a focus group interview with the staff and statistics on the participants. The general conclusion from the study is that the configuration of the balance between individual needs and the regulation of the group facilitates the participant’s motivation. In the discussion, the importance of institutional complexities upon developmental pathways for the participants will be emphasized.

Keywords: Group based interventions; Motivation
The destructive potential of evaluation and how to avoid it

The destructive potential of evaluation: and how to avoid it

U. Schiefer1, A. Carvalho2, L. Bal-Doebel3
1Iscte Iul Avenida das Forcas Armadas, Sociology, Lisboa, Portugal
2Cea Iscte Iul Avenida das Forcas Armadas, Sociology, Lisboa, Portugal
3EP, Romania, Bukarest, Romania

At the beginning of the workshop participants will discuss the theme and in a participative way visualize their ideas and design and produce an installation on one of the walls of the conference centre which is accessible to all conference participants.

These will be encouraged to also write up their experience on coloured sheets of paper and fix it to the wall and to comment the already exposed papers. The goal is to identify ways in which evaluations may have negative and even destructive effects.

The design of the installation will be decided by the workshop participants – proposed headings are: Where is the destructive potential of evaluation? How does it operate? What are its causes? What are the destructive impacts of evaluation?

Evaluation happens in fields where different, often conflicting interests meet – and very often in intercultural contexts which add complexity and increase risks. The “underbelly” of evaluation has not yet received the attention of the evaluation research.

Evaluation is a relatively new field of expertise and knowledge production, as well as a relatively new area of economic activity, which is conquering an ever increasing base of revenue. It is in strong competition with other more established fields (management, quality, research, etc.). Therefore practitioners and researchers are focusing on promoting its benefits and ignoring and disclaiming the more negative aspects. It might be useful to reframe this participative symposium: all stakeholders in evaluation should know the destructive potential in order to avoid it, and learn how to do so.

All stages of evaluation count: what to evaluate, who is involved in this decision, terms of reference, selection of evaluators, evaluation design, research, report writing, presentation and communication of results, utilization and, last but not least, payment and other sanctions.

Keywords: Destructive potential of evaluation; Negative impact; Competition for revenue base
Monitoring an evaluation in developing countries

M&E in the health sector in Rwanda

L. Inberg, N. Holvoet
Institute of Development Policy and Management, University of Antwerp, Antwerp, Belgium

One of the key principles in the reform agenda set for donors and recipients (Paris Declaration 2005, Accra Agenda for Action 2008) is ‘management for development results’, for which one indicator is formulated: ‘number of countries with transparent and monitorable performance assessment frameworks to assess progress against (a) the national development strategies and (b) sector programmes’ (OECD/DAC, 2005: 10). Despite commitments made by donors and recipients, progress in the implementation of reforms in this area is slow. The latest update of the Comprehensive Development Framework (CDF) report (World Bank, 2007), on which indicator 11 is based, reveals that only three out of 54 countries surveyed had result-oriented frameworks that were deemed adequate (OECD/DAC, 2008a:58–59). Even though most countries do have a number of monitoring and evaluation (M&E) activities and systems in place, there is often a lack of coordination and harmonization between them. For a nationally owned and properly functioning performance assessment framework, however, an appropriate organisational of a national M&E system with clear division of responsibilities between different levels and layers of government, with clearly identified information streams and accountability structures between central and line ministries, between local and national is crucial. This article will focus on the sectoral M&E arrangements development in the context of the health Sector Wide Approach (SWAp) in Rwanda and their relationship towards the national M&E system. The health sector M&E system will be assessed on selected criteria of policy, methodology, organisation, participation of non-state actors, capacity and quality. Specific attention will be paid to the place of Joint (Sector) Reviews within the M&E system and the role of sector working groups.


S. F. Nkwwenke, V. Fargeon, C. Offredi
University of Grenoble, Department of Economics, Grenoble cedex 9, France

Different methods have been explored to measure health inequalities as a key for evaluating health policies. These have included the distribution of capabilities as a way to assess its moral dimension. In developing economies, especially in African countries, a number of investigations have highlighted the fact that the increase in health expenditure will only contribute to reduce health poverty gap if such expenditure is efficient and access to health services becomes more equitable. This paper contributes to the study of health equity in developing countries, with particular attention paid to child and maternal health services. We examine disparities of health and health care between rural and urban areas, regions, and population subgroups in Cameroon, where health policies were redirected in the 90’s with focus on improvement of equity in health and health care.

Based on data from Demographic and Health Surveys, health policies outcomes were assessed through the relationship between socio-economic conditions and health inequalities. The contribution of various subgroups (from poorest – Q1 – to better-offs – Q5) to health inequalities was appraised in terms of health service utilization, access to medication and participation. We found that these policies have to some extent had a positive impact on immunization (effectiveness), while impact on health care utilization and reduction of financial burden of diseases for the whole population, especially the poor, was negative (inequity). In such a context, improvements in several dimensions of governance would make access to health services appreciably more equitable. Compared to expected outcomes from the policies, lots of challenges in term of policy implications are identified as ways to improve design and implementation of equity-oriented health policies.

Keywords: Health policies; Capability approach; Equity; Concentration indices; Cameroon
Evaluating institutional versus individual capacity development in Afghanistan: lessons, challenges and innovation

R. Guillard, M. Marin
International Development Law Organization, Evaluation Unit, Roma, Italy

Background
The International Development Law Organization (IDLO) has been working in Afghanistan since 2001. During the past nine years, IDLO has implemented a number of legal reform programs focused on supporting the National Justice Sector Strategy (NJSS) of Afghanistan by building the capacity of justice and legal institutions through legal training and technical assistance initiatives. The program stakeholders include the Office of the Attorney General; the Ministry of Justice; the Legal Aid Organization of Afghanistan; and the Independent National Legal Training Center. In recent years, the IDLO Afghanistan program has had three evaluations. The focus has been finding effective ways to determine the results of external efforts to build the capacity of Afghan legal professionals, practitioners and institutions to better deliver much needed legal services in Afghanistan.

Objectives
1) The presentation/discussion will share some of the methodological approaches and key lessons of these evaluations; 2) explore some of the challenges to evaluating institutional capacity development initiatives in a vulnerable/fragile state like Afghanistan; and 3) share some of the innovative methods that were employed to measure how individual and institutional capacity develops, and the relationship between the two.

Methods
Methods include a large survey of legal professionals throughout Afghanistan, and evaluating against a capacity development model developed by the European Centre for Development Policy Management (ECDPM, 2008).

Conclusion/Application to Practice
Based on these evaluation experiences, IDLO will update its model for capacity development as a guide to better understanding how external actors can effectively support institutional improvements as part of legal reform. They will also add to our knowledge of evaluation methods which are effective for measuring capacity development, particularly in conflict/post-conflict settings.

Keywords: Innovation
Evaluating local development

Outcome indicators to measure the effectiveness of local economic development interventions of South African local authorities

B. Rabie¹, F. Cloete¹
¹University of Stellenbosch, School of Public Management and Planning, Cape Town, South Africa

Governments increasingly focus on the outcomes of their programmes and policies, necessitating the development of outcome based monitoring and evaluation indicators. Promoting the economic welfare of citizens is a critical objective of local government, but the absence of specific indicators for LED measurement hampers their ability to successfully determine whether their efforts are achieving the expected results.

This doctorate research produced a systematic framework of concrete contextual output and outcome indicators to enable and encourage the measurement of the results of the various local economic development interventions adopted by local authorities. A review of LED-related literature produced a comprehensive list of potential LED interventions. These interventions were categorised into business development, locality improvement, community development, or improved governance interventions.

A list of indicators was compiled for each intervention by harvesting existing indicators from various international sources. While some indicators measure local economic development, many indicators with a more general focus had to be adapted to the local economic development context. These indicators were used as departure point for the development of more context specific output and outcome indicators for each of the identified LED interventions. The final LED indicator framework was circulated to selected M&E and LED scholars and practitioners for comments which were then used to further refine the indicator framework.

These results may be used to compare alternative LED interventions and to identify the most promising market and community driven LED interventions. This can improve strategic decisions to maximise the limited available resources for LED for the greatest positive economic and social impact of good local governance.

Keywords: Outcome evaluation; Outcome indicators

Evaluation of Social Investment Funds in developing countries in the context of decentralization

H. Voigt, E. Terberger, R. Rischke, M. Sieburger
KfW Entwicklungsverband, FZ E, Frankfurt am Main, Germany

Since the end of the 1980s decentralization as part of a broad movement towards local democracy and accountability of administrative entities vis-à-vis civil society has become a major concept of development co-operation. In this context, Social Investment Funds (SIFs) evolved to be one of the main tools to support decentralization via Financial Cooperation, although these type of funds originally were designed to quickly deliver small scale local infrastructure in settings in which there was a need for quick relief. Accordingly, SIFs objectives changed from a pure infrastructure implementation scheme, working parallel and in replacement of government structures, to an instrument with the double mission of fostering small scale infrastructure investments as well as strengthening local administrative authorities, participation and accountability in a decentralization setting. Whether this shift in objectives was adequately reflected in project design and in the design of (rapid appraisal) ex post evaluations of SIFs is the topic of a cross-cutting analysis of 46 ex post evaluation reports of SIF projects in 14 countries carried out by KfW evaluation department.

The meta analysis reveals that SIFs were rather effective in the provision of (small) municipal or local infrastructure; but less so in fostering long run sustainable operation and maintenance. Contrary to SIFs double mission, the evaluation reports concentrated on the subject of infrastructure. When trying to fill this evaluation gap by collecting at least some evidence on SIFs contribution to strengthening decentralized administration and governance, the results are sobering.

Decentralization processes in the sample countries seem to have dwindled over time from an ambitious start to stagnation. SIF projects in no way were able to change the course of politics, and in most cases even failed in strengthening local administration. New designs of Community Development Funds promise better results. Evaluation methodology needs to adapt to new generation SIFs.

Keywords: Decentralization; Social Investment Fund
Performance measurement systems in pilot municipalities in Angola

E. Sarmento
Universidade de Aveiro, Degei, Lisboa

The need to give priority to M&E has thus become widely accepted by most developing countries Governments. But in many cases, national
monitoring systems have been designed to meet donor data requirements and end up focusing on budget aspects of resources spent on
national priorities and on national progress towards the MDGs.

What has been missing is the intermediate layer, that is performance information on the intervening steps in the results chain, involving
Government activities, outputs and services provided, and their outcomes and evaluative evidence linking Government actions to actual results.

Since 2002, the Government of Angola has initiated reforms. It has embarked on decentralization and strengthening of local governance.
Phase II of UNDP Decentralization and Local Governance (DGL) was designed with strong participation of provinces and municipalities.
Building on the guidelines of UNDP support for DGL, a monitoring and evaluation system at the municipal level started to take root, by
mid 2009. The pilot projects were launched in municipalities of 5 provinces of Angola. The need to systematize budgeting and
programming, with planning and monitoring, have led to the development of four integrated main database systems, called SIM (Integrated
Municipal System of databases).

Adding up to the challenges of measuring such an overreaching policy such as decentralization, comes the pre-requisites of designing,
buidling and sustaining a results-based monitoring and evaluation system at a municipal level, one that is also able to interact at both the
sector, provincial and central level.

This article is an account of this process and on the need to build on the momentum of the work done so far by the DGL at the local
Government level, to engage on to further efforts to promote M&E systems in Angola.

Keywords: Monitoring; Performance systems; Municipalities; Angola; Evaluation
Evaluating agricultural and rural development programmes

The Mix of Methods: towards a framework for anticipating validity threats in evaluations of agricultural value chain support

G. Ton, M. De Ruyter De Wildt
Wageningen UR, LEI, Wageningen, Netherlands

Impact evaluations of value chain interventions are challenging: outcome indicators are often multidimensional, impact is generated in dynamic and open systems, and social embeddedness is constraining external validity. Therefore, there is a strong case for theory-based evaluation, where logic models indicate how the intervention is expected to influence the incentives for people’s behaviour. The key assumptions inherent in these casual models can be tested through the observation and measurement of specific outcome indicators using mixed methods in triangulation.

The paper presents a framework to evaluate the design of these mixed methods to assess change and impacts in value chain configurations, using the four types of validity threats to the evaluative conclusion. We apply the framework on three impact assessment experiences, we have been involved in, and reflect on the feasibility to improve its methodological design: micro-irrigation technology supply in Nepal and Zambia; farmer field schools for sustainable tea production in Kenya; business service development for horticulture in Bangladesh.

For each case we identify the major evaluative conclusions that has been derived from these evaluation studies, revisit the evidence-base that supports it and recommend the use of complementary methods that may be included to strengthen the research design. In doing so, we use a structured approach, expected to develop in a framework to be used in internal peer-review.

Keywords: Methods; Development; Credibility; Research; Impact

Catalyzing learning, change, and increased effectiveness of development funding through Impact Evaluation: the case of Sasakawa for Africa

R. La Rovere¹, D. Areedo²
¹ CIMMYT, CIMMYT, Addis Ababa Ethiopia, Italy
² Addis Ababa University, Agricultural economics, Addis Ababa Ethiopia, Italy

The Sasakawa Africa Association (SAA), a NGO active in agriculture and extension funded by the Nippon Foundation (NF), is known to have had successes in Africa over the past 2 decades, and is often cited as a model for promoting productivity growth. Yet SAA is also seen as an organization that adopts an old fashioned model of development that may not always meet recent challenges, and ‘unable to learn’. SAA also used to lack a systematic way to document its impact objectively. Alike other donors under pressure to be transparent on the effectiveness of development funding, the NF commissioned to CIMMYT an innovative, evidence-based external monitoring and assessment system (MIE project, sg2000ia.cimmyt.org) for the 2006–2010 investment period. The paper describes the challenges faced by the evaluator teams and partners in evaluating SAA, and how the methodological and management challenges were addressed. We summarize the lessons we learned in running the MIE within the daily challenges typical of Africa. After 4 years of panel data collection, several changes catalyzed by the MIE – yet not automatically or fully attributable to it – have taken place: now SAA is better targeting those who are really in need, discontinued ineffective activities, and focused in fewer areas its previously thinly spread efforts. Also, the existence of the MIE led SAA to replicate it by institutionalizing an ‘internal’ M&E system for its new phase (2011–13), also to keep attracting those wanting to invest in agriculture in Africa. Yet many challenges remain. The paper uses data from the MIE that demonstrates the link between more effective extension services and increased adoption and yields, allowing SAA to move into evidence-based policy advocacy. It also uses the results of a recent external study that critically reflects on the effectiveness of the MIE itself in meeting its evaluation goals.

Keywords: Agriculture; M&E; Learning and change; Africa; Development effectiveness

R. Conner¹, E. Gotor², R. La Rovere³, A. Aw-Hassan⁴, H. Waddington⁵

¹Social Ecology – UCI, Policy Planning and Design, Irvine, USA
²Bioversity, Cgiar, Rome, Italy
³Cimmyt, Cgiar, Mexico City, Mexico
⁴ICARDA, Cgiar, Aleppo, Syria
⁵International Initiative for Impact Evaluation, Synthetic Reviews Office, London, United Kingdom

There is much discussion in the international evaluation community about impact evaluation (IE), particularly in evaluating the effects of development projects in developing and transitional economies. IE advocates a shift away from monitoring and evaluation that focuses primarily on inputs, processes, outputs and outcomes and a shift to impacts, both short- and long-term. There have been new institutes and initiatives such as 3ie that have generated significant new funds to support impact work. Unknown to many evaluators, however, the agriculture research sector has had a long history of impact work known as impact assessment (IA). The Consultative Group on International Agricultural Research (CGIAR) has a special focus on IA. CGIAR involves 15 centers working in developing countries and regions around the world. Its mission is “to achieve sustainable food security and reduce poverty in developing countries through scientific research and research-related activities in the fields of agriculture, forestry, fisheries, policy and environment.” CGIAR’s impact work focuses on measuring its success in achieving reductions in poverty and hunger, improvements in human health and nutrition, and enhancements in ecosystem resilience. To measure these impacts, CGIAR centers undertake ex post impact assessments (epIA). These epIA are different from what most in the evaluation community call “impact evaluation” (IE) in their focus, scope, scale and methodology. Whereas IE usually focuses on a specific small-scale project and involves longitudinal pre- and post-measurements and a counterfactual (such as non-intervention control or comparison groups), epIA occurs 5–10 years after a project and focuses back to document large-scale impacts, usually with a counterfactual using baseline measurements. This session will describe the general epIA approach, illustrate it with several studies, and then contrast it with IE known to most in the international evaluation community, via a discussant familiar with IE and IA.

Keywords: Developing economies and regions; Impact assessment; Impact evaluation; Ex post impact assessment; Agriculture sector
Evaluation of management practices

Evaluating the support of ecological management practices in Viennese enterprises: Lessons learnt from the evaluation of the programme “EcoBusinessPlan Vienna”

I. Fischl¹, S. Mayer¹, R. Schüle²
¹Austrian Institute for SME Research, Evaluation, Vienna, Austria
²Wuppertal Institute for Climate Environment and Energy, Energie-Verkehrs- und Klimapolitik, Wuppertal, Germany

The “EcoBusinessPlan Vienna” is a support programme of the Municipal Department 22 – Environmental Protection Vienna that is implemented since 1998. The rationale of the programme is to help enterprises, by funded consulting, generating profits through ecological management practices that benefit both the environment and the enterprises, ensuring high quality and financial benefits in each company.

The accompanying evaluation of the 3rd programme period (2006–2009), conducted by the Austrian Institute of SME Research in cooperation with the Wuppertal Institute for Climate, Environment and Energy, included an analysis of the design, the implementation as well as the effects and efficiency of the programme. The evaluation team applied both qualitative and quantitative approaches in a cyclic manner including feedback loops to the client over a time period of three years to meet the main requirements of an accompanying and systemic evaluation.

This paper primarily deals with the results and conclusions for the future strategic development of the programme design and processes, i.e. it focuses on inconsistencies between the level of (strategic) goals and instruments and the key role of the environmental consultants within the programme context. Furthermore, the paper presents the essential conditions of success for the future development of the programme identified, based on the manifold results of the different surveys and analyses, hereunder: consulting demands of the enterprises, aspects of mobilisation, motivation for participating in the programme, PR-aspects etc. Due to annual surveys – covering different relevant players in the context of the programme – the evaluation is able to show the different perspectives of the consultants and the consulted enterprises, contrasted to each other and moreover provides longitudinal comparisons. Apart from that, the lessons learnt from evaluating a programme supporting consulting as an instrument to accelerate the implementation of measures by enterprises for achieving reduced ecological damage will be presented.

Keywords: Accompanying Evaluation; Ecological management practices; Funded environmental consulting

Norwegian-Russian performance audit on fisheries management

Y. H. Arnesen
The Office of the Auditor General of Norway, Performance Audit Department II, Oslo, Norway

Background
Norway and the Russian Federation cooperate on the management of fisheries in the Barents Sea and the Norwegian Sea pursuant to international agreements. In the mid 2000s illegal overfishing represented a challenge to the sustainability of the fisheries.

Motivated mainly by this issue, the Office of the Auditor General of Norway and the Accounts Chamber of the Russian Federation are cooperating on conducting performance audits on the management of the shared fish stocks. Because of its bilateral nature the fisheries management is well suited for joint audits. A parallel audit was completed in 2007, and follow-up investigations are ongoing.

Objective
The main objective of the investigations is to assess goal achievement in implementing bilateral fisheries agreements.

Methods
The first parallel audit had common audit questions and audit criteria, but the audit operations were conducted independently and separate reports were written. A document with common assessments was prepared on the basis of the reports. This way of reporting proved to be successful because it granted the supreme audit institutions autonomy while at the same time enabling them to communicate a common statement to the parliaments and the public.

In the follow-up investigation some of the interviews with government agencies are conducted in cooperation. Observations of fishing vessel inspections were also conducted. Data is partly analysed in cooperation. The goal is again to complete separate reports and a document with joint assessments.

Results and conclusions
The cooperation makes the Norwegian and Russian auditors better equipped to assess the fisheries management in both countries. When the audit results are reported to the parliaments, the legislature will thus be provided with a better basis for its policy decisions.

Furthermore, the collaborative audit process has in itself contributed to management improvements and more openness on a controversial issue.

Keywords: International cooperation; Performance audit; Fisheries management; Norway and Russia
Do chief executive officers and middle managers of community health centers identify the same skill sets middle managers need?

P. Walsh
Eastern Michigan University, School of Health Sciences, Michigan/Ypsilanti, USA

Federally qualified community health centers (FQHCs) in the United States are a safety net for over 20 million individuals with limited financial resources. The FQHCs are primarily nongovernmental organizations. Since their inception in the 1960s, there have been studies evaluating FQHCs’ clinical and cost-effectiveness. There are virtually no studies regarding the competencies that middle managers need to have to be effective in carrying out the mission of the community health centers: to provide primary health care to the medically uninsured and underinsured.

The primary objective of this exploratory study was to identify the knowledge and skills middle managers working in FQHCs need to be effective in their roles. The study method used a mail survey in June 2009 to 400 chief executive officers and middle managers working in FQHCs. Both groups were asked to complete the survey based on their perceptions of the knowledge and skills middle managers need to be effective.

The CEOs and the middle managers were asked to rate the importance of 34 activities that were listed under the categories of planning, recruitment/selection of staff, supervision/management, quality management, communication, and outreach/advocacy. Under each section, managers were asked to identify, from a list, the top three skills or knowledge that they need to perform the above-mentioned activities.

Preliminary results indicate that middle managers identified leadership, oral communication and knowledge as the three most important skills that they need to be effective managers. The three least important were cultural competency, creativity, and monitoring. The CEOs responses are in the process of being analyzed and will be included in this paper. The responses for both groups will be compared to identify their similarities and differences.

The FQHCs will be able to use this information to evaluate their programs and to identify training needs of the managers.

Keywords: Competencies; Managers; Healthcare; NGO
Why gender and human rights matter for evaluation? 
Evaluation from a gender perspective. Sponsored by UNIFEM (part of UNWOMEN)

“Why gender and human rights matter for evaluation?”
B. Sanz Luque, S. Reddy, F. Etta
1 UNIFEM, Evaluation Unit, New York, USA
2 Agden, Afrea, New York, USA

This panel emerges from the need to debate on the implications of gender perspective for evaluation, based on the hypothesis that the type of change that equality policies and programmes intend to influence imply not only shifts in the normative order, but also forms of social, economic and cultural transformation that evaluation should aim to capture.

Furthermore, it builds on the assumption that all public policies, even if not targeting gender inequality and other forms of inequality directly, do have differential effects – intended or not intended – on different groups of people who may be discriminated against, calling for evaluation to advance in analyzing these dimensions beyond data disaggregation by sex or other variables, which would result in increased quality and credibility of the evaluation practice as a whole.

The panel is approached from a double perspective: first, it will look at the particularities of evaluating programmes and policies whose main goal is gender equality and women’s rights, as well as other forms of equality from a human rights approach; secondly, it will promote an analysis on the meaning of applying a gender and human rights perspective to evaluation.

Evaluation of gender equality programmes/policies:
• What do we measure when we evaluate gender equality programmes or policies?
• What approaches and methods work best? What are some promising new approaches/methods?
• How can such evaluations be used to improve gender policies?

Evaluation from a gender perspective of non-specific gender equality programs:
Why is a gender perspective relevant for evaluation? Do evaluations fail when they do not take into account the gender dimension?
What does it mean to do an evaluation from a gender perspective? Is Evaluation from a gender perspective a matter of process or substance?
How to consider other inequalities/forms of discrimination and a human rights approach in evaluation?

Keywords: Gender; Human rights; Evaluation

Evaluation from a gender perspective
M. Bustelo, B. Sanz, S. Reddy, P. Alvarez, J. Espinosa, F. Etta, K. Hay, A. Callerstig
1 Universidad Complutense de Madrid Facultad de CC. Políticas y Sociología, Dpto. C. Política y de la Admón. II., Pozuelo de Alarcon (Madrid), Spain
2 UNIFEM, Evaluation Unit, New York, USA
3 Independent consultant and Complutense University, Master in Evaluation Of Public Policies and Programmes, Madrid, Spain
4 Complutense University, PhD program in political science (Public policies evaluation), Madrid, Spain
6 IDRC, Delhi, India
7 Linköping university, Tema Genus, Department of Thematic Studies, Linköping, Sweden

The objective of this symposium is to explore gender-sensitive evaluations as well as the meaning and possibilities of evaluating with a gender perspective. As part of the policy making process, and following the aim of the internationally recognized gender mainstreaming strategy, evaluation is an important part to be conducted under a gender – and human rights and other inequalities – perspective; however, the implications of what it means to do evaluation from a gender perspective still remain to be thoughtfully explored.

In this symposium five papers are presented. Bustelo will present a positional paper which aims at clarifying the various ways in which evaluation and gender interplays. Sanz & Reddy will address the evolution of gender equality and human rights responsive evaluation in UNIFEM and the broader UN, share experiences and progress made to date, and highlight the remaining challenges and areas that require further thinking and exploration. Alvarez explores how the gender dimension is not yet integrated into the evaluation criteria and procedures, discussing the opportunities and challenges to include gender equality within evaluation criteria by both engendering standard evaluation criteria as well as adding alternative criteria to the existing ones. Espinosa presents different evaluation experiences in analyzing gender issues in international development evaluation describing the main challenges of carrying out gender-sensitive evaluations. Finally, Etta chronicles the attempt to engender the Paris Declaration in Africa with the support of UNIFEM, and briefly presents the indicators and draws out the attendant politics of the effort made.

These five papers have already been submitted as individual papers as well.

Keywords: Gender and evaluation; Gender mainstreaming; Evaluation from a gender perspective; Gender policies & programmes
Evaluation and Employability

Learning through ongoing evaluation for innovation and job creation

S. Jansson¹, G. Brulin², L. Svensson³, C. Ehneström³
¹ The Swedish ESF Council, Evaluation and Monitoring, Stockholm, Sweden
² Swedish Agency for Economic and Regional Growth, Audit and Review, Stockholm, Sweden
³ APeL FoU, Lindesberg, Sweden

In Sweden a national evaluation report concluded that the short term investments made in the former EU program period had not made any impression whatsoever on Sweden’s regional economic development. During this third program period Sweden will dispose more than € 3.1 billion from the EU Structural Funds. The commission has recommended a new approach to evaluation – ongoing evaluation. The goal is to create prerequisites for better sustainability and greater impact by continuous learning in the projects and regions and to ensure that the experiences made from the projects are communicated.

To implement this new evaluation approach in Sweden a number of activities are organized – a textbook about learning evaluation has been produced and has been used in a number of university courses. Networks for evaluators are arranged around the country in order to exchange knowledge and experiences. Four different programs for support to the projects are organized on a national level among one is about learning and strategic impact. There are also five thematic groups with a mission to analyze, systemize and disseminate knowledge.

In all the role of learning and ongoing evaluation is very important, both in the projects and on the program level. Our research findings points to the usefulness of an on-going evaluation approach. But in practice it has been difficult to implement. Often the evaluators become too involved in the projects and act more like projects leader than critical evaluators. The focus in the evaluation is often on the short term outcomes, not the long term effects. A too small proportion of the project funds seem to be used for evaluation which might lead to quality reduction. In the paper and seminar we will discuss how the attempts in Sweden to introduce an on-going evaluation in the EU-programs can be strengthened.

Keywords: Ongoing evaluation; EU Structural Funds; Sustainability; Continuous learning; Long term effects

Training programs evaluation: reflections concerning the evaluation of a policy for employability

M. Terraseca, A. S. Costa
Faculty of Psychology and Education Science, CIE-Centro de Investigação e Intervenção Educativas, Porto, Portugal

We intend to discuss the implications that the choice of an evaluation model has in the production of evaluation results and the contributions that evaluation can actually make to improve and redefine sectoral interventions.

This reflection is based in an evaluation of a concrete intervention that intended to achieve social and territorial cohesion by favouring territorial intervention in areas of employment and training, improving employability, developing new skills in emerging sectors and encouraging institutional articulation in the territory. The intervention strategy was to implement vocational training programs on a large scale, covering a heterogeneous population of an administrative area of Portugal and a variety of training areas, in order to increase the potential of individual's employability.

The evaluation team could have opted for:

– a "classical model of evaluation", in which it was crucial the comparison between obtained results and outlined objectives. In this case, the object of evaluation would be pre-defined: the training program and its results.

– an evaluation model that would allow questioning the objectives and fundamentals of the intervention itself, seeking to understand the adopted strategy adequacy concerning the intervention wider goals. In this case, the evaluation would allow thinking the problem that the intervention was intended to solve through a broader perspective and a complex thinking by considering new relationships between training, qualifications and employability. Therefore, allowing to uncover implicit modes of thinking those relationships.

The option by this second path allowed the problematization of intervention decisions and to increase an innovative evaluation model that enable the evaluation of complex systems by considering their own complexity.

We defend that this evaluation model is particularly relevant to the evaluation of complex systems – as training systems – because it provides the fundamentals of an evaluation process able to question the actions in order to enable their own transformation.

Keywords: Employability and training; Evaluation models; Complex systems
Evaluation of tendering procedures in the field of public employment services in Flanders

S. Devisscher, D. Sanders
IDEA Consult, Labour Market and Socio-Economic Policy, Brussels, Belgium

In Flanders, policy makers are moving since the early years 2000 towards an all-inclusive approach for the integration of jobseekers into the labour market. Under influence of different factors, the Flemish public employment service introduced tendering procedures and consequently competition in the market of private employment services. A first pilot project was launched in 2005 in which 6000 pathways for long term unemployed jobseekers were allocated to private employment service providers, operating on both a non-profit and profit basis.

The functioning and the results of the pilot project were externally monitored and evaluated by IDEA Consult. The evaluation started the moment the private service providers were selected and continued until nearly all jobseekers ended their pathway. Several research methods were used: literature review, descriptive quantitative analysis and multivariate statistical analysis, multiple interview rounds with the different stakeholders involved, and large scale surveys.

The quantitative analysis was well developed thanks to the availability of detailed client monitoring data of the PES. The data included a range of characteristics of the job seekers, information concerning the intake, the duration of the reintegration programme, the intensity and type of training, and the employment results on short and longer term. Moreover, results from the other research methods were introduced in the data set to fill in information gaps. For example, the interviews and the surveys, gave a view on the content and the intensity of the counseling.

This data set provided the basis for extensive monitoring on the approach towards job seekers and results of the private partners and for a regression model to identify the success factors. Next to that a control group was set up ex-post to provide a benchmark for evaluating the employment results of the pilot, using a propensity score matching methodology.

Keywords: Employment; Job seekers; Public employment service; Tendering procedures
Contribution Analysis: Evaluating impacts with rigour and pragmatism

Contribution Analysis: Evaluating impacts with rigour and pragmatism

J. Toulemonde¹, E. Wimbush², S. Montague³, S. Lemire⁴

¹ Eureval and Lyon University, Partner, Lyon, France
² Edinburgh University and NHS Health Scotland, ESRC Placement Fellow, Edinburgh, United Kingdom
³ to be confirmed Performance Management Network, Partner, Montreal, Canada
⁴ Ramboll Management, Copenhagen, Denmark

Contribution analysis is a recent, rigorous and highly feasible tool for evaluating impacts. It involves a step-by-step test of the programme theory through systematic, evidence-based, and contradictory falsification efforts. This technique is recommended in a number of evaluation guidelines, but it has not yet been applied at large scale. The purpose of this session is to describe the current state of the art in this area.

The session includes three presentations reporting on three contrasted practices:

Steve Montague reports on the introduction of contribution analysis in the Canadian Cancer Society which is the largest not-for-profit group in Canada, the country where the contribution analysis was born ten years ago.

Erica Wimbush reports on the findings from her current work to assess whether and how contribution analysis might be applied to local partnership planning, evaluation and performance reporting. She builds upon two case studies of contribution analysis applied to local partnerships in Scotland: one is seeking to address inequalities in mental well-being in a defined community and the other is developing a strategy for tackling anti-social behaviour.

Jacques Toulemonde reports on three EU evaluations where contribution analysis was applied. All addressed complex cause-and-effect issue, including that of cost-effective achievement of complex impacts. These evaluations had to assess the impacts of three policy instruments, i.e. the open method of coordination of EU Member States, the EU decentralised agencies, and a sector budget support in the area of development aid.

The authors and the participants will discuss the issues of (1) practicalities and feasibility, (2) generalisation and transferability of lessons learned, (3) areas of relevance.

Those who wish to learn more may also attend the pre-conference workshop devoted to this tool.

Keywords: Impact; Cause-and-effect analysis; Contribution analysis
Evaluation of public sectors institutions, policies and programmes

Towards a more efficient and effective government? Some experiences and considerations from Flanders

D. Verlet¹, M. Vandersmissen², B. Van Hoorebeek¹, L. Vandeputte¹, B. Van den Haute¹, E. Vermoesen⁵, E. Groené⁵, G. De Schepper⁵, S. Spiessen⁵, J. Matthys⁵, F. van Laere⁵

¹ Flemish government, Research Center of the Flemish government, Brussels, Belgium
² Flemish government, Internal Audit of the Flemish Administration, Brussels, Belgium
³ Flemish government, Department of Education and Training, Brussels, Belgium
⁴ Flemish government, Public Personnel Agency, Brussels, Belgium
⁵ Flemish government, Public Governance Department, Brussels, Belgium

Although policy evaluation has always been important, today there is a rising attention for policy evaluation in the public sector. Especially there is the search for good governance. Good governance rises in a diversity of contexts and is not a new phenomenon. However, in the wake of different debates, like the one on efficiency and effectiveness in the public sector, it is noted as a bright spot in troubled discussions. At the end, we can not be against a good governance. In this framework, efficiency and effectiveness of the public sector became a central concern of different governments and stakeholders. However, efficiency and effectiveness are only two principles of good governance, besides other principles, for example concerning ethical conduct, accountability, transparency, integrity, stewardship, etc. cetera.

In this panel we would like to organise a dialogue between academics and policy makers from several policy and research fields, at different policy levels. The discussion will be based on the several papers we present in this panel. We deal with some methodological issues and the normative debate on the public sector. Besides the theoretical positioning, special attention is given to the way good governance is incorporated within the reforms planned and –although partly– implemented within the Flemish government.

Among others, in this panel we try to explain and analyse, from a broad and interdisciplinary perspective, the growing importance of output and outcome measurement in the public sector. How to measure public policy performance, taken into consideration the specific nature of the public sector? What are the underlying normative positions? How can governments build capacity in order to increase government-wide performance, or in order to be able to monitor the performance of their policy-programs?

Keywords: Efficiency; Effectiveness; Good governance; Government performance
Ethical process

Evaluation and accountability: so close, so far
E. Garcia Sanchez
Universidad Carlos II De Madrid

Accountability is often considered one of the main uses of evaluation. But, are evaluation and accountability actually so close? If we explore the dimensions and characteristics of accountability that several authors have pointed out, we conclude that the relationship between evaluation and accountability is not as clear and simple as it seems at first glance.

In this paper we will review the main studies both on accountability (Manin, Przeworski y Stokes, 1999; O’Donnell, 1997, 1998, 1999; Schedler, Diamond y Plattner, 1999) and on evaluations uses (Caracelli y Preskill, 2000; Weiss, 1990, 1999; Forss, Rebien y Carlsson, 2002; Patton, 1997, 1998, 2000, etc), in order to explore the theoretical elements on which the links between evaluation and accountability are built. Our aim is to analyse how evaluation can play a key role as tool of political accountability

Keywords: Accountability; Evaluation; Politics

The control of the programme manager to the consultant influence the ethical standard and transparency in programme evaluation
A. Hanif
Concern Worldwide, Bangladesh, Learning & Sharing

Background
the national and international organization in developing countries measure the impact of the programme carrying out the evaluation recruiting the consultant by the programme managers. the consultant carry out the study following a terms of references, but questions is what extend the consultant get freedom reporting the real findings of the evaluation? there is carried out a study to understand “the control of the programme manager over the consultant and influence the ethical standard and transparency in programme evaluation” in developing countries.

Objective
to generate knowledge about “the control of the programme manager over the consultant and influence the ethical standard and transparency in programme evaluation”

Methods the study followed a qualitative methods applying some key informant interviews and Focus group discussions with the national and international consultant, some programme managers, others relevant programme saffs and review the documents.

Results
the consultant shared that some of the programme managers are very much concerned about accountability to their donors and their higher management and their tendency to capture as their favorite results. the programme managers try to channelize more resources with highlighting the positive impacts. they never thought about the downward accountability to the programme people. the programme manager control the consultant in favour of designing study methods, tools and writing report as targeted impact of the programme.

In addition, the programme managers shared that the consultants have capacity on single methods instead of using mixed methods, they give less concentration designing the methods and applying the tools. the financial benefits also a factor to use the less rigorous tools and they follow the instruction of the programme managers instead of the terms of references. the ethics and transparency here is questionable due to their lack of capacities.

Conclusion:
the freedom consultancy ensure real impact and transparency of programme evaluation.

Keywords: Control; Influence; Mixed methods; Accountability to donors; Downward accountability
Transparency for Evaluators and Evaluations

J. Greene 1, H. Simons 2

1 University of Illinois, Educational Psychology, Illinois Champaign, USA
2 University of Southampton, Professor of Education and Evaluation School of Education Faculty of Law Arts and Social Sciences, London, United Kingdom

Evaluation in government is increasingly valued as a tool of new public management (NPM) interests, often advanced to serve such ideals as results-based government accountability and credible evidence for decision making, as well as contributing strongly to government transparency. Yet the assumptions underlying these aspirations are not always transparent, in particular, how these NPM aspirations address the full spectrum of stakeholder interests in a democracy. In order for evaluation to genuinely and meaningfully serve a democratic plurality of interests – in a new public management context or any other – commissioners and evaluators must be substantially more transparent than they currently are about their own commitments, values, and stances and how these are present in their evaluation practice.

This argument is grounded in the premise that all commissioners of evaluation and evaluators bring certain predispositions, assumptions, values, and stances to their work. These include beliefs about what kind of evaluation generates the most warranted and useful results, assumptions about social realities and their meanings, favorite substantive theories and constructs, and stances related to the proper role of the evaluator and what constitutes ethical conduct in the field.

Yet, too often these assumptions and stances are not acknowledged as a routine part of evaluation practice. The values being espoused are most evident in the criteria chosen to make judgments of evaluation quality. But in too many evaluations these criteria are not made explicit; nor are they subjected to diverse stakeholder input and dialogue. So, the assumptions, values, and stances that underpin evaluation intent and practice remain hidden – masked equally well by our avowed independence and impartiality or our championship of learning or social justice as underlying rationales for evaluation; and by government’s appeal to methodologies that promise conclusive evidence to inform policy.

We will present this argument, in part, as a narrative.

Keywords: Transparency, Values
Participants and co-production V

Democratising evaluation: A case study of local essential services
D. Walker
Monash University, PSL, Melbourne, Australia

The provision of local basic public services such as primary healthcare and education is regarded as essential for reducing poverty in developing country contexts. Improving a range of public services in a given community can create a “virtuous circle” of positive social change and improved child and community well-being. However, reforming such services “from above” is often deeply problematic, and approaches designed to bring services closer to citizens such as decentralisation are often disappointing.

This paper outlines and analyses processes of monitoring and evaluation of service provider performance by impoverished local citizenries in developing countries, employing a rights-based approach known as Citizen Voice and Action (CV&A). When otherwise marginalised actors both monitor and evaluate local public services, questions arise about an enabling environment for these activities, about their status as rights-bearing citizens, their capabilities to be producers of knowledge based on their experiences as service users, their empowerment to be decision-makers about service reform, and whether this can effect lasting reform. In CV&A, citizens decide what they value about local services as the basis of taking action for reform, they collectively generate their own data and knowledge, and own and utilise them to make decisions which seek to deepen accountability and pursue valued human rights.

This paper will employ a case study of CV&A from Brazil or Uganda, and will draw on various documents including examples of data generated by citizens, and evaluation and project reports. It will touch on ethics and transparency issues, and draw on Sen’s capability approach. It examines how enabling conditions cognisant of local politics and addressing power imbalances are created and sustained, and the challenges in doing so. Lastly what implications does CV&A as a participatory evaluation approach have for evaluation utilisation, policy and planning, and its own evaluation.

**Keywords:** Service provision; User evaluation; Social accountability; Human rights; Democratic evaluation

Developing a differentiated and complementary approach to external evaluation
C. Mutch
Education Review Office, Corporate Office, Wellington, New Zealand

As evaluation has evolved as a discipline and become embedded in a range of government, public and private spheres, the notion of public interest has become more contested. The purposes, philosophical underpinnings, processes and intended outcomes all carry with them notions of whose interests are being served. If the purpose is accountability – then accountability for what and to whom? If the purpose is improvement, what role does evaluation play in who determines and drives the improvements? If the evaluation is externally driven or internally motivated, how are the different interests of stakeholders taken into account?

This paper, set in the context of a system of national educational evaluation in New Zealand, describes how the answers to these questions have changed in response to external trends, especially from the US, UK, and Europe, and internal factors, particularly a call for more contextually and culturally relevant practices. The Education Review Office (ERO) is the government agency charged with the external evaluation of all pre-tertiary education institutions (i.e., early childhood, primary, and secondary; public and private; English medium and indigenous) in the New Zealand education system. ERO provides the government with information on the quality of education in individual institutions and the system as a whole. This role has expanded over time to respond to the broader range of stakeholders with legitimate interest in the evaluation findings – schools, school trustees, parents, community, the Ministry of Education, the government minister responsible for education, professional groups, cultural and ethnic groups, and other government agencies. ERO has taken on an evaluation capacity building role in an attempt to make the process more of an evaluative dialogue to meet the needs and interests of all stakeholders. This has led to the development of a differentiated methodology and the complementary approach to evaluation described in this paper.

**Keywords:** External evaluation; Evaluation partnerships; Evaluation stakeholders; Evaluation capacity building; Participatory evaluation
Participatory monitoring for improved sustainability

M. Körner

The road to rural drinking water supply projects in developing and transition economies is often lined with broken pumps, pipes and storage systems and frequented by women and children tracking water from distant or polluted sources. Sustainability remains low.

The objective of this paper is to share experience how a participatory monitoring mechanism with self-monitoring and planning tools contributed to improved sustainability of community owned rural water supply installations (sub-projects) rehabilitated and constructed with support of the Rural Water Supply and Sanitation Project in Kyrgyzstan (RWSSP).

The project attached high priority to monitoring sustainability of the sub-projects. Conventional approaches and methods did not meet the specific requirements of the project but their elements were used to design a participatory monitoring mechanism with self-monitoring tool – a sustainability check list- at its core to extract information on likely sustainability from the communities which were involved in shaping the mechanism. By maintaining an ongoing dialogue and providing feedback to the communities, the mechanism also contributed to their empowerment.

The sustainability check list includes indicators and scores for sustainability factors that could be influenced by the project and communities: Economic, organizational and cooperation with partners. Scores indicate the level of risk and need for action. The check list offers: easy-to-use and understandable structure; application in modules allowing focus on critical factors; numeric scores to enter in Management Information System (MIS) for planning and reporting; it can be completed within 1 hour including discussion and consensus seeking.

Use of the mechanism demonstrated how monitoring can in addition to generating information on sustainability also contribute to its strengthening by fostering the link between monitoring and interventions as well as by increasing local awareness of risk factors and capacity to identify and to address them. Transferring ownership to the local levels contributed to their empowerment and changed perception. Monitoring was no longer perceived as an examination imposed by outsiders. It was recognized that there is no need to pretend better results because realistic results help to see the problems and look for solutions; self-monitoring itself motivated to actions and to improved sustainability.

Likelihood of sustainability has been assessed before, during and after construction.
Evaluation in a range of educational settings

Benchmarking and evaluation in higher education

I. Barbulescu, A. Miroiu
National School of Political and Administrative Studies, Political Science, Bucharest, Romania

The paper discusses the use of benchmarking in the evaluation of higher education institutions. Quality assurance procedures developed by specialized agencies across Europe include the appeal to benchmarking, as a powerful yet flexible method, based on comparison. It is widely agreed that it can be used to formulate solutions to organisational problems, to highlight strengths, weaknesses of higher education institutions and help improve performance measurement.

The paper presents a case study. It focuses on the approach developed by the Romanian Agency for Quality Assurance in Higher Education (RAQAHE) in the past four years. Although limited to this case, the paper intends to highlight a more general conclusion concerning the trends in higher education benchmarking.

The paper is organized as follows. In the first section the main principles of benchmarking in higher education are discussed. Instruments, scope and objective, and values of the benchmarking process are presented. In the second section we analyse the initial approach to benchmarking developed by RAQAHE and analyse the impact of the results of external evaluations of about 40 universities on the benchmarking process. The third section presents some trends in the benchmarking process: the move from an arm-chair to a more empirically oriented approach; the clear classification of the indicators used for benchmarking into four classes: input, process, output and outcome indicators; and an increased role of the quantitative analysis. We give examples of the way in which the use of evaluation indicators of RAQAHE are challenged by this new approach to benchmarking.

More generally, we argue that benchmarking requires a continuing and systemic practice. Usually it is difficult to empirically identify “the best practice” in absolute terms; therefore, in order for benchmarking to work it is seldom better to use “good” practices as referential.

Keywords: Benchmarking; Quality assurance
Constructing an index of educational effectiveness: the visit to Genoa Aquarium

D. C. Torrigiani
University of Genoa, D. S. A., Genova, Italy

This paper illustrates an evaluative research process – committed by the Genoa Aquarium and developed by the Department of Anthropological Sciences – aiming to construct an index of educational effectiveness of the visit, which allowed the company management to get a synthetic indication of effectiveness and to identify more easily the strengths and weaknesses of the path from the educational point of view. To do this, after a preliminary analysis of the bibliography of reference, the first step was the conceptual clarification of the evaluation object, sought through focus groups and interviews in which managers and operators of the structure were invited to more properly define the educational objectives of the visit. Then questionnaires were constructed and filled in by the visitors before the visit, to ensure that their educational expectations on the visit coincide with the objectives identified by company management, and after the visit, to evaluate its effectiveness. In the final phase the staff has been involved in the indicators weighting, allowing to re-evaluate the scores obtained on the different indicators with the administration of questionnaires and to calculate, on a 0-100 scale, an index of educational effectiveness that respects the priorities defined by managers and experts of Genoa Aquarium. The main index strength, thus calculated, is that it reflects the priorities assigned by company management, and on this basis gives a synthetic feedback on the educational effectiveness of the visit. Its intrinsic limit, is that one can not assume the overlap between the priorities assigned to the objectives by managers and operators, on one hand, and the relative priority of user expectations on the other, and this is an important element of uncertainty for the definition of corporate strategies.

Keywords: Educational effectiveness; Evaluative indicators; Evaluative index; Mixed methods; Participatory evaluation
Evaluating trade industry and investments

Impact evaluation of UNIDO interventions in Sri Lanka in the area of Trade Capacity Building

P. Loewe
United Nations Industrial Development Organization, Evaluation Group, Vienna, Austria

The capacity of developing countries to participate in global trade is dependent on well functioning National Quality Systems (NQS). The share of NQS projects in UNIDO’s technical assistance portfolio is growing but evaluating the impact of systemic capacity building interventions such as the ones related to strengthening NQS is methodologically challenging.

This impact evaluation was conducted by the UNIDO Evaluation Group with funding from NORAD and in cooperation with the Institute for Policy Studies in Sri Lanka. It covered a series of UNIDO interventions in Sri Lanka aiming at upgrading public sector laboratories in the areas of metrology, product and process testing and focused on the impact of these interventions on three major export sectors (garment, tea and fisheries).

The evaluation found that the improvement of laboratory testing and calibration services has contributed to improving Sri Lanka’s export performance, in particular for fish exports. However, attempts to attribute these improvements to the UNIDO interventions would be flawed by a multitude of external factors, which for obvious reasons cannot be eliminated through “control groups”.

The evaluation applied the so called “theory based” approach and has led to valuable insight into the specificities of the causal chains (or rather causal “fabrics”) in the area of Trade Capacity Building in general and Standards, Metrology, Testing and Quality (SMTQ) in particular. Potential impact drivers were identified and possibilities explored for better targeting future SMTQ projects towards poverty alleviation.

Keywords: Impact evaluation; Trade Capacity Building; Theory based evaluation approach; United Nations Industrial Development Organization

Investment Funds in emerging markets – How to evaluate them?

I. Yong Protzel
European Investment Bank

For several years the EU has supported the development of the private sector in developing countries as a key element of its poverty reduction strategy. The European Investment Bank (EIB), the financial arm of the EU, has been requested to participate to this effort and carry out, i.a., risk capital operations.

The EIB has recently undertaken the evaluation of its investment funds operations in the Southern Mediterranean (FEMIP) and African-Caribbean-Pacific (ACP) countries. These operations are financed with EC and Member States’ resources. The aim was to understand what has worked (or not) and why and to draw lessons for future operations. Several challenges were faced:

• It was the first attempt for the EIB to evaluate investment funds in emerging markets
• It covered two geographical regions for which the EIB has different mandates (sets of objectives, governance rules, etc.)
• These are operations carried out through financial intermediaries whose objectives are different (but complementary) to those of a development institution
• There was a lack of meaningful benchmarks for many of the indicators
• Investment funds have a usual term of 12 years, therefore the sample was older than for other evaluations
• Private equity is an activity that has significantly evolved during the period covered by the evaluation.

The paper will outline the scope and methodological approach used for this evaluation. In particular, it will present the evaluation grid developed for evaluating individual investment funds. The grid includes specific indicators to measure development effects. The paper will also present findings and conclusions. Finally, methodological lessons will be drawn from this evaluation.

Keywords: Investment funds; Evaluation methods; Defects; Emerging markets
Evaluation issues in educational context

Quality oriented research; conversations about assessment in schools or building dynamics into both design and results

A. Reinertsen, A. Reinertsen
Nord-Trøndelag University College, Education, Steinkjer, Norway

My work on evaluation is both part and product of an ongoing research project designed for and directed towards developing Sustainable Assessment Cultures in Norwegian Schools. Student assessment, adapted teaching and school as a learning organization constitute the basic theoretical framework. Poststructuralist and critical feminist research traditions’ stance of language however, the background for how these theories is “understood” and “applied”.

The goal is the development of a more nuanced education language for use in assessment. And as far as research is concerned, the development of a more nuanced language for use in research. On the one hand, we try together with students and teachers to create and provide systematic and well – grounded support for teachers in making their judgments and to recognize the different ways students learn, and the different ways in which their learning is expressed and can be forwarded. On the other hand and as far as the students are concerned the goal is to create environments in which the critical discernment of quality becomes the key aspect of learning.

Cross curricular or multi theoretical and interdisciplinary approaches are important and both means and goals. Methods used involve open observation in classrooms, regular semi-structured interviews with teachers, students and leaders, group interviews, document studies and filming of activities of particular interest.

Four conversations about assessment have so far been developed at “Cityhill” High School. These are conversations not primarily about what assessment is, but what it does.

a) Education as experience and taking risks.
b) Conceptual certainty and how it can be assessed?
c) Written formative assessment on an electronic platform
d) Assessing practical exams with a focus on the oral aspect.

Ultimately this might be conceived of as a way of building (professional) autonomy through subjectivity and experiential learning; a way of writing competence forward.

Keywords: Qualitative Research; School Reform; Writing; Subjectivity; Applied Ethics

Measuring professional uncertainty of teachers

Jacob Gujski
Kibbutzim College of Education Technology

This study focuses on “professional-uncertainty feeling of teachers” (PU) as perceived by teachers from different academic subjects in high-school and college. It was aimed to conduct an empirical construct validation of a questionnaire for measuring PU, based on the approach of Jackson (1986) and McDonald (1992) with elements taken from Light (1980) and Fox (1957) who dealt with professional-uncertainty in the medical-profession. The questionnaire included 14 items with 3 subscales and 5 multiple-choice items, with an open-ended option. The multiple choice items deal with different strategies applied by teachers to acquire information about their students that might reduce his or her PU. A factor-analysis (Varimax rotation with Kaiser Normalization) was performed on the 14 items and three factors matching the three theoretical factors have been revealed: feelings during the teaching process, feelings at the end of the teaching process, and general awareness of uncertainty in teaching.

Reliability of the three scales was computed using Cronbach’s alpha: 0.80 for the first factor, 0.78 for the second and 0.61 for the third. Removing one problematic item resulted a growth to 0.62, which means that this item had no significant influence on the scale in spite of the findings from factor-analysis.

Considering the small number of items in each scale and the small sample of (44) we can conclude that the internal consistency of the three scales is quite high.

The answers to the multiple-choice items regarding the strategies applied to acquire information about the students’ knowledge during the teaching – learning process and its end were similar.

The most significant finding of this study was that PU, as represented by 14 items, has construct validity as revealed by factor analysis. As a result the questionnaire can be used to measure PU in future studies.

Keywords: Professional uncertainty
“When leaders make the difference” – The leadership of school’s organisational learning through teacher research groups

A. Ovdvin
Sogn og Fjordane University College

The presentation will outline some important research findings within the field of educational leadership on quality improvement and school development in Norway. Organisational learning is the main theoretical foundation. The findings are built on a qualitative study. The methodological approach is frame analysis of dialogues on quality in a specific lower secondary school in Norway, and the design is built on observations and interviews.

The paper will discuss the challenges a school is facing when the school's aim is to be a learning organisation. The research question is: What are the strategic interventions for the school's leadership to enhance organisational learning through quality improvement?

The school’s quality improvement is formulated in the following goals:

• Creating a learning culture in the entire organisation.
• Quality analysis carried out systematic in the organisation, and reflection as a crucial part of the analysis.
• Quality work as a part of the Curriculum planning. Defining areas for development, plan for developing steps.
• Making the teachers qualified – systematic professional upgrading. Connected to the developing areas for the school.

Teachers in this school are organised in school based research groups, and these groups are giving direction for the school's development. The core issue of this paper is clarification of the discourse of leadership and execution of pedagogical leadership in this process on quality improvement and processes of change. The paper will discuss analysis of focused situations where the leaders change directions of the developing process. An important finding is that the leaders’ values and styles are crucial for leadership that enhances learning.

**Keywords:** Evaluation; Leadership in schools; Learning organisation
Circular Dialogue and other dialectical methods of inquiry

R. Hummelbrunner
OEAR Regionalberatung, Graz, Austria

Background
Circular Dialogues are systemic forms of communication among people with different perspectives. They use the perspectives and views of participants as a resource e.g. for appraising/validating different experiences or identifying joint solutions. Circular Dialogue is a systemic technique rooted in the tradition of role plays, which is increasingly being used in organisational development, but is little known in the evaluation field. They are an example of a family of methods, which use using dialectics in dealing with multiple perspectives and opposing viewpoints in order to gain meaning.

Objectives
The purpose of the paper is to describe Circular Dialogue and three other examples of dialectic methods of inquiry. Thus evaluators will be shown how this type of methods function, which could be particularly useful when faced with different perspectives or interpretations.

Methods
In addition to Circular Dialogue, three other dialectical methods will be briefly outlined, namely: Convergent Interviewing, Contradiction analysis and Option One-And-A-Half. Participants will be familiarized with their main elements, but are also made aware of pitfalls to avoid (i.e. defensive reactions, trivial use of “straw dolls”). All the methods are careful in their processes to scaffold acceptance rather than rejection of contrary evidence and debate, as well as avoiding contrariness just for the sake of it.

Conclusion/Application to practice
The main benefit of dialectical methods of inquiry is challenging mental models, “common sense” assumptions and established beliefs. They help to answer the following questions, which are often posed in evaluations:

• What are the different ways in which people see or can see a situation?
• What are the exceptions or contradictions to the way in which people see or can see a situation?
• How can exploring and making sense of these differences enhance our understanding of a situation?

Keywords: Dialectic methods; Multiple perspectives; Differences in perspectives; Dialogue

Evaluation in education: partnership, guiding or guarding?

B. Walczak1, G. Mazurkiewicz2
1Warsaw University, Institute of Applied Social Sciences, Warszawa, Poland
2Jagiellonian University, Institute of Public Affairs, Krakow, Poland

This presentation is a result of the reflection over implementation of the new governmental regulations concerning evaluation of the formal educational system in Poland. We would like to focus on a role of legally embedded evaluation in “programming” the education. We are in particular interested in transforming culture of evaluation from evaluation fulfilling diagnostic function to evaluation forming dialogical function. There is a threat, that the values connected with dialogical evaluation may be easily deformed into the figure of the guide (“I will show you where to go”) or even to always-present guard from Bentham’s panopticon (“You will know before I open my mouth”).

Indeed, the system of evaluation in education, just implemented in Poland consists of several, broad expectations that form a general vision of Polish educational system. But during the training with schools’ directors, we often hear “beautiful vision, but please let me see your research tools, and my school will be ready”. It clearly shows the urgent need of mediation between the dialogue, partnership and self-dependency in education and top-down determined expectation.

We would like to show some ideological roots (and also outcomes) of this approach to evaluation in education, then try to argue that this ideology won’t assure school autonomy without authentic changes within the field (among professionals) that will make self-regulation and self-consciousness possible.

Keywords: System of evaluation in education

The evaluation of the Austrian initiative programme “Laura Bassi Centres of Expertise” and its policy implications

A. Doerflinger1, S. Mayer1, E. Heckl2
1Austrian Institute for SME Research, Evaluation Innovation and Technology, Vienna, Austria
2Austrian Institute for SME Research, Evaluation Employment and labour market Entrepreneurship, Vienna, Austria

The proposed paper will be based on the findings from the current evaluation of a new Austrian RTDI funding programme “Laura Bassi Centres of Expertise” (LBC) launched in 2009 by the Austrian Federal Ministry of Economy, Family and Youth. This initiative programme aims at equal opportunities for women in cooperative research, at establishing a new research culture, at offering new career options for all research personnel at LBCs. The call for proposals was open to all topics in the areas of natural science, technology and life sciences at the gateway between science and economy. Currently, eight LBCs in cooperation of science and industry are funded, all of them headed by female researchers.

The initiative programme defines itself as a “learning initiative”. It tried to adopt new approaches to contribute to the above mentioned goals. The accompanying evaluation the paper is derived from analyses these new approaches in the light of the programme goals. The methodology will in general rely on qualitative approaches: interviews, focus groups and workshops targeting and will be complemented with a standardized survey, addressing heads and staff of the LBCs.
The evaluation has a strong focus on learning elements that include not only learning within the evaluated funding scheme itself but deliberately aims at the evaluation and identification of programme elements that could be transferred to other RTDI-measures. In this sense, the initiative programme could also be conceived as a “teaching” programme. Involved international jury experts already consider it as European best-practice. Hence, depending on the findings, the information gained from the programme and from the evaluation could also provide other research programmes with impetus for a more gender-equal design and increase the gender competence of the Austrian scientific community.

**Keywords:** Accompanying evaluation; Initiative RTDI funding programme; Gender; New research culture; Policy learning

---

**P-04 Sustainability analysis of an evaluation policy: the case of the primary health care in Brazil**

**E. Felisberto 1, E. Freese 2, L. C. Albuquerque Bezerra 1, C. K. Almeida Alves 1, I. Samico 1, Z. Hartz 1**

1 Instituto de Medicina Integral Prof Fernando Figueira – IMP Research Department – Research Group of Health Evaluation, RECIFE-PE, Brazil

2 Centro de Pesquisas Ageu Magalhães/Fundação Oswaldo Cruz – CPqAM/Fiocruz, Public Health Department, RECIFE-PE, Brazil

Since the year 2003, the Ministry of Health of Brazil develops the National Policy for Monitoring and Evaluation of Primary Health Care. It is composed by six components integrated of various specific projects towards the institutionalization of evaluation in the Brazilian health system. This study presents the application of a methodology of sustainability analysis of this Policy from the period of 2003 to 2008 taking into account the identification and categorization of critical events representing its institutionalization process. It is an evaluative research with two units of analysis: Federal Management of the Primary Health Care and the Health State Secretariats (HSS). It is a multiple-case study and data were collected by interviews and documentary analysis. It was used the Critical Incidents Technique. The events were analytically categorized as of memory, adaptation, values and rules and temporally classified as specific of the implementation, of the sustainability and mixed, being or not favorable towards sustainability. The Federal Management and one HSS were classified in the medium degree of sustainability and the other HSS as a strong degree. The results indicate concomitant events and suggest to be pondered since those of adaptation of activities, adequacy and stability of resources demonstrated a great influence on the others. Innovations and Technical Capacity Development are considered the most important results towards the sustainability.

**Keywords:** Sustainability; Institutionalization; Program evaluation; Health evaluation; Evaluation

---

**P-05 A strategic evaluation of “Nanotechnology, multifunctional materials and new production processes” (6th Framework Programme): Measuring what cannot be measured?**

**S. Ruhland 1, I. Fischl 1, B. Romanow 2, M. Gustafsson 2**

1 Austrian Institute for SME Research, Evaluation InnoTech, Vienna, Austria

2 Oxford Research AS, Evaluation, Kristiansand, Norway

In 2009, the Austrian Institute for SME Research (KMFA) and Oxford Research Norway were awarded by the European Commission with conducting the strategic ex-post evaluation of the thematic priority NMP. While currently being in the final phase of the project, this paper deals with the particular challenges that came with the little word ‘strategic’.

This strategic evaluation has been labelled by the Commission as an ‘experiment’. The terms of reference indicated what the Commission was hoping for, yet most of the evaluation questions raised were not different from other evaluations, leaving us puzzling over how to connect questions like “What was the nature and value of results produced?” with the idea of evaluating at the strategic level. Ultimately, we successfully developed an approach that set out to gather information at project level (but not numbers of patents, publications etc.) combined with extensive literature analysis and a comprehensive set of interviews with experts. However, the main task was to identify and eventually operationalise the objectives as included in the relevant strategic documents. Therefore, we had to extract the respective objectives, categorise them and develop an idea how to measure the impact NMP FP6 had without completely relying on qualitative, non-representative information. Furthermore, an understanding had to be achieved if NMP FP6 was more than the sum of its parts (projects). In the end, we had to find a way to measure what in fact cannot be directly measured by combining experts’ qualitative assessment with a survey among project coordinators based on the idea of approximation (i.e. that the sum of contributions of individual projects provides insights into the overall impact of NMP FP6). The whole exercise as for its experimental nature resulted in a mutual learning process for the evaluation team and the European Commission.

**Keywords:** Strategic evaluation; Framework Programme; Ex-post evaluation; Nanotechnologies

---

**P-06 Evaluation and Power**

**U. Schiefer**

Instituto Universitário de Lisboa, School of Sociology and Public Policies, Lisboa, Portugal

Evaluation has come a long way since it was claimed as to be “speaking truth to power” or to be an exercise for the “empowerment of the disempowered”. Most, if not all evaluations happen in spaces where power is a more or less important dimension of relations between different actors and stakeholders. Usually evaluation occurs at nodes in space and time where different power-lines intersect. The different projections of power into the “evaluation space” naturally influence evaluation in all its phases, from its inception, to the design, the research, and the transformation of research results, the presentation and communication of evaluation findings and to the use of the evaluation results by different stakeholders. Evaluations are influenced by power – but inversely they also bear influence on power structures, projections and relations as well as on the way power is organised and perceived. Evaluators and other stakeholder do in practice know this. They adapt their behaviour to their perception of the power fields they are working in, be these organisational, political, military, security, economic, religious, magic, personal, intellectual, cultural or other forms of power.
According to the different ways power is produced, applied and felt in different cultures and societies, however, evaluators apply different strategies; not always with much success when dealing with unfamiliar forms of power. Strangely, evaluation methodology seems to have been designed around the power phenomena; in most evaluation approaches, theoretical and practical, power in all its types, forms and dimensions seems to be non-existent. This paper will try to fill in parts of this blank space and produce theoretical, methodological and practical knowledge about power that could be included in evaluations in all its phases.

**Keywords:** Power production; Power projection; Types of power; Methodology and power

### An impact evaluation of a displaced community after twenty years – Case Study of Pantabangan in the Philippines

**S. Tamondong Giraud**
AfDB Temporary Relocation Agency (Tunis) African Development Bank Group, OIVP, Tunis, Tunisia

**Background**
Development induced displacement (DID) is a current global issue, where communities are forced to leave their homes and land for public purposes. It is associated with land acquisition for development. Examples are construction of dams, mining, military installations, airports, industrial plants, railways, roads and even sanctuaries/forest reserves. DID is a social problem affecting more and more people as countries industrialize. An estimated ten million people annually enter the cycle of forced displacement in two sectors alone: dam construction and urban transportation. In the pursuit of better living conditions, governments continue to construct dams and these are co-funded by international banks providing loans in the name of development. The PANTABANGAN Project was an example when the former Philippine President Marcos dreamt to make his country show case of an industrialized state in Asia and embarked to build Pantabangan Dam in the late 60s. The Project inundated Pantabangan town, displaced 13,000 people and 7 outlying villages.

**Objectives**
The impact study examined what happened to the displaced community after twenty years—analyzed differential impact on various groups of population in the community composed of farmers, fishermen, local merchants and government employees. Who were the winners and the losers in this process? How did they emerge? How did women cope in this process?

**Methods**
Methodologies employed mixed methods, quantitative as well as qualitative. Community participated in the evaluation in assessing their past and current lives.

**Results**
The evaluation study found out that, during displacement crisis, when men lost traditional roles/jobs associated with land, women emerged as leaders and breadwinners. After twenty years, the impact of involuntary displacement created an emancipatory/economic benefit among women.

**Conclusion/Application to Practice**
Why promote development when it creates violation of the right of another? Findings of this research demonstrate development policy implications and practice, particularly those targeting women.

**Keywords:** Pantabangan; Philippines; Case Study

### React to the crisis in a multiannual programming framework

**R. Cagliero, F. Pierangeli**
INEA – National Institute for Agricultural Economics, National Rural Network, Rome, Italy

The ongoing economic crisis has triggered significant phenomena of change, which should lead to a review of policies oriented to increase farms competitiveness within a changing scenario. Alongside the effects of the crisis on a global scale, other major events occurred: chain shocks, significant changes in European policies (Health Check of the Common Agricultural Policy, the reform of Common Market Organisations), and some modifications in the regulatory framework at the local level.

However, the process of review should not be undertaken chaotically, but fit coherently in the logical framework of the programming approach, ensuring consistency with the intervention logics of rural development, derived from the regulation.

Taking into account the stiffness of the programming framework, which characterises the rural development policy, it was foreseen an element of dynamism and flexibility within the evaluation processes. In this sense, the ongoing evaluation is called to support the process of editing programs, to rationalize their own instruments and, above all, to demonstrate a strong dynamism.

Therefore, by observing the tangible reactions to modifications of the context, the purpose of the paper is to provide operationally a critical remark on choices undertaken and affecting the support for farms competitiveness.

The question arising is twofold. How have the priorities – coherently expressed within the logical framework of the Rural Development Programme (RDP) – found concrete implementation? and have these priorities changed over time because of the global crisis and other modifications?

Following these issues how can the evaluation advise – at the same time – choices consistent with a well-framed programme (i.e.: RDP) and suitable to address changes occurred over time?

**Keywords:** Rural development; Logical framework; Crisis; Tender; Modernisation
**P-09**

**An analysis study on suitability of new national R&D budget overall coordination system**

J. M. Shim¹, Y. I. Park², Gu-Sun Park³

¹ Korea Institute of S&T Evaluation and Planning, HRST Policy Division, Seoul, South-Korea
² Ewha Womans University, Ewha council for R&D innovation, Seoul, South-Korea
³ KISTEP (S&T policy division)

Based on the concept of the overall coordination of Korea's national R&D program, this paper means to look into the status of current overall coordination system and suggest its rectification. For this, this paper is composed as follows. First, it examines R&D system and governance theory with precede theory and study related R&D. Then, it renewes the concept and scope of R&D Program and searches examples of advanced country such as United States and Japan. Second, dividing periods from the before launch of NSTC(National Science and Technology Council) to the overall coordination of Lee Myung-bak's administration, this paper looks into its change and analyze its characteristics and achievements. Lastly, based on consequences from in-depth analysis of stakeholders, this paper examines change of Korea’s R&D overall coordination totally and suggests its rectification.

**Keywords:** National R&D budget

---

**P-10**

**Evaluation for enlightened policies? A Swedish case study on employment policies.**

A. Sohlmän

Sohlmän Senior Consultants, Head, Stockholm, Sweden

**Background**

This study is a case study examining the utilisation of evaluations of employment training in the national political process in Sweden. How does the Government’s interests in evidence based policy (EBP) influence its utilisation of evaluations?

**Objectives**

One could argue that in an EBP-context for a sector dominated by a rather technocratic discourse, an instrumental use of evaluations would prevail. The fact that research and evaluations have a high standing in labour market policy more or less forces the government to recur to evaluations. But this can be done more to the profit of the Government if the evaluations are used not only instrumentally but also to legitimise its policy. My hypothesis therefore is that I should look for evidence of both instrumental and symbolic use of evaluations by the Government.

**Methods**

This is an analytic case study. My main data sources are communications from the government. I have also had personal contacts with representatives from the Ministry of Labour and I made a few interviews and was myself investigating employment training 2006-2008.

**Results**

The investigation shows that the Government has used evaluations differently depending on whether they fall within policy areas that the Government is interested in or not. Evaluations of the first kind are used instrumentally but evaluations of the second kind are not. Employment training happened to be one area the Government was politically not interested in so e.g. the evaluations showing positive results for employment training and possibilities to further improve it were dismissed as unscientific and relegated to the “analytical framework” for future use.

**Conclusions**

The findings of this investigation i.a. raises the question as to the responsibilities of evaluators to mediate between policy makers and those on whose behalf policy makers are expected to work.

**Keywords:** Utilisation; Employment policies

---

**P-11**

**Participatory Development of a Logical Framework in Brazil: The Labor Gymnastics Experience**

L. Vilarins¹, M. Fontes², R. Crivelaro², F. Pereira¹

¹ Social Services of Industry – Brazil, Unitep, Brasília, Brazil
² John Snow Brazil Consultancy, Research and Social Impact Evaluation, Brasília, Brazil

**Background**

Labor inactivity in the workplace is one of the main sources of occupational health disorders. According to the International Labor Organization, the long hours of inactivity contributes to high rates of absenteeism in many countries.

**Objective**

The Brazilian Industrial Social Services Organization (SESI) is the largest third-sector organization, sponsored by the private sector in Brazil. In 2001, SESI developed a Labor Gymnastics program in the workplace. Trained professionals were contracted to promote daily fifteen minutes physical activity sessions in the workplace. In 2009, the program has been implemented in all 27 states of Brazil, reaching over 500,000 workers in the industrial and service sectors. The objective of this study is to demonstrate impact evaluation quality using a participatory evaluation process among labor gymnastics program staff.

**Methods**

Using a logical framework matrix, program's staff from all states of Brazil and of different hierarchy levels (mediated by an evaluation team), established consensus about programs values, and also set objectives, goals and indicators aiming a follow-up impact evaluation. This participatory approach was key for better defining the goals of the program and its impact indicators.
Results
A participatory evaluation process engages program's staff with the results and can lead to the effective use of the findings, communication of the outcomes and a continuing rethinking of the program. 30 consensual efficacy and over 50 impact indicators were established and used in the development of evaluative research tools. Staff members also used the final logical framework as a tool for better defining the redesign of the program.

Conclusions
Participatory evaluation processes produce not just research tools, but they empower those involved in the development of a social program. However, some trade-offs of a participatory impact evaluation approach vs. independent/external evaluations should also be considered.

Keywords: Logical Framework; Impact; Labor Gymnastics

Implications of top down and bottom up evaluation of multilevel governance

A. Hanberger
UCER, Umeå University

The purpose of this paper is to discuss what it means to evaluate implementation of public policy from a top down and bottom up approach. Implementation of local safety and security policy, a case of multi level governance, is used as an example. Actors and institutions from the public and civil sphere, and different levels of government constitute this policy field. Whether a top down or bottom up approach is applied is mainly determined by the evaluation questions, and evaluators' methodological preferences. However, the choice of approach is also a choice of what role evaluation can play in multi-level governance. A bottom up approach is more inclusive and invites stakeholders to engage, not only in the evaluation process, but also in governance and implementation. The paper first describes how implementation of local safety and security policy can be studied from a top down and bottom up approach. Next, two cities' local safety and security policy, including implementation of the Swedish Safety Protection Act, are analysed from a top down and bottom up approach. Thirdly, the paper discusses implications of the two approaches for multilevel governance and democracy.

Keywords: Multi-level governance; Top down and bottom up evaluation; Implications; Local safety and security policy

Up-scaling and Out-scaling: Monitoring & Evaluating the Paris Declaration, Gender Equality and the Africa Gender & Development Evaluators Network

F. Etta
African Evaluation Association, Accra, Ghana

Women's empowerment and gender equality have been on the agenda of international development for an undeniably long time yet full victory cannot still be claimed. This paper tells the story of the political journey and its implications of the Africa Gender and Development Evaluators Network (AGDEDN).

With support from UNIFEM a small band of evaluators located in Africa attempts to take on and engender one of the most significant political documents of contemporary development practice. The challenge is huge but not without its moments. The first ever set of gendered indicators to mirror and correspond to the official monitoring indicators of the Paris Declaration complete with an accompanying monitoring framework were developed.

This paper chronicles this attempt, briefly presents the indicators and draws out the attendant politics of the effort to engender the Paris Declaration.

Keywords: Paris Declaration; Gender Equality; Indicators; Monitoring & Evaluation

The complex dynamics of citizen participation: Engaging minority groups

P. de Lancer Julnes
University of Baltimore, School of Public Affairs, Baltimore, USA

Consistent with the theme of this year's EES Conference, “Evaluation in the Public Interest: Participation, Politics and Policy,” my presentation will deal with the question of who gets included? How? Does it work and why? In this era of accountability for results, there’s been a chief in the role that citizens can and should play in the process of governance. Some see citizens as important players in shaping the quality and responsiveness of government programs in their community (Symonds, 2005; Petts, 2001). Others believe that citizen participation is the cornerstone of democracy (Symonds, 2005; Petts, 2001). My presentation, based on a series of interviews and survey data will specifically focus on the role that minorities (people of different ethnic and racial backgrounds of the ruling majority) play in the process of governance in the United States. I will discuss whether traditional assumptions about participation hold for these groups, the methods that are used to engage them and the potential for the success of the various approaches. The goal is to provide practical advice for evaluators and policy makers interested in involving citizens of diverse backgrounds.

- A Presentation made at the American Society for Public Administration Center For Accountability and Performance Symposium's Leadership of Results-Oriented Management in Government. Washington, DC. February 11–12.

Keywords: Citizen participation; Minority population; Engagement; Democratic Governance
P-16 Beyond results: theory-based evaluation of a syringe exchange program in Catalan penitentiaries (Spain)

R. Crespo, M. Pla, N. Codern, A. Farré, A. Cardona
Autonomous University from Barcelona, Chair of Qualitative Research, Barcelona, Spain

Background
Birckmayer and Weiss (2000 p. 409) describe different advantages associated with the use of theory-based evaluation models. Specifically, they suggest its usefulness in terms of “evaluation planning”, “program design” and “knowledge creation”.

The evaluation of the Syringe Exchange Program (SEP) in Catalan penitentiaries (Spain), driven by the departments of health and justice of the Catalan Government, has been a collaborative “theory based evaluation”, which emphasizes some considerations regarding the experience proposed by Birckmayer and Weiss.

The SEP is a harm reduction program targeted to drug users in prison to avoid unsafe practices (share the syringes and use them more than once), prevent infectious diseases, and promote the acquisition of healthy behaviors and habits. The program provides venipuncture kits to prisoners who request it, as long as they systematically return the syringe previously provided.

Objectives & Methods
The results show that theory-based models facilitate the scheduling of an evaluation process that focuses on the objectives achievement while exploring the “black boxes” that condition the objective compliance.

However, given the complex nature of the program, the evaluation of the SEP also suggested some challenges that must be considered: The first is related to the convenience to merge the sequential gaze of the “theory of the change”, with the recursive and emergent logic that pervades complex programs such as the SEP. The second, related to the desirability of going beyond the theory-based models advantages previously described, in terms of planning assessment. Specifically, with the need to convert “theory of change” in a device that facilitate a critical analysis of the frameworks and intervention models, where evaluated programs are built on.

Keywords: Theory based evaluation; Complex program evaluation; Frame analysis approach; Syringe exchange program evaluation

P-17 Citizen approach on regional and global level evaluation?

H. Isola-Miettinen
National Audit Office, Strategic Unit, Helsinki, Finland

Evaluation is usually understood to be the activity, which takes place on national level and inside the national public administration organisation. The first general problem raised up in this paper is the question concerning citizens: should we give some more active role to citizens in the public sector evaluation processes? Should we legitimate evaluation processes through the citizen approach?

One lifts up in this paper the citizen approach problem not only on national level but on regional level and global level. Is citizen approach needed on these levels? Namely, generally taken, the evaluation has legitimating effect; it is giving the feedback information to decision-making. Feedback about the activities how the “politics” have been transformed into “policies” and finally, for example, into the services, like judicial services or other services like health, education services etc.

States are based on certain political system, usually we talk about democratic decision making where national Parliaments make the politically important top decisions. EU as regional actor is political system, not a state. It is not always very clear what actors make the political top decisions in EU. The most problematic question is how we define the global actors and global criteria in this respect? There are global actors who regulating global issues. Do they have responsibility in this respect? Do they have such political responsibility that is interesting and important question in respect of evaluation and citizen approach? This paper focuses the possibilities of citizen approach in the public sector evaluation on all those levels, especially on regional and global level where the regulations and convention consist from more open principles than rules. The paper focuses and reflects could the human right principles be appropriate criteria in making evaluations from citizen approach perspective.

Keywords: Regional and global evaluation; Citizen approach

P-18 Implementation of evaluation in Polish education system. Opportunities and challenges.

T. Kasprzak, A. Borek
Jagiellonian University, Institute of Public Affairs, Kraków, Poland

We would like to focus on the presentation of the process of internal and external evaluation implementation in schools and institution in the whole Poland, in the framework of the pedagogical supervision reform (project carried out since 2009 by the Ministry of Education, in partnership with Jagiellonian University, in the Framework of 3rd Priority of the Human Capital Operation Program, Action 3.1, co-financed by the European Social Fund).

Although it is not the first contact of Polish schools with evaluation, this process is of a nation-wide character and has a legal basis (resulting in creation of a new function in pedagogical supervision system – Evaluation Inspector). At the moment the processes of a pilot program, checking research tools and training of evaluators (a couple of hundreds people), school head teachers and institution directors (3600 people) take place.

We wish to present basic dilemmas of ideological (values and goals of evaluation), methodological (research methods, methods of evaluators and evaluation recipients trainings) and implementation (usage of the evaluation results to strengthen schools and educational system) character. These considerations will be illustrated by the results of the first thirty three evaluations carried out in Polish schools and institutions.

Keywords: Evaluation in schools and institution
The “evaluation planning process” of Spanish National Strategies in Health: the case of National Strategy of Rare Diseases

R. Crespo, M. Pla, N. Codern, A. Cardona, A. Farré
Autonomous University from Barcelona, Chair of Qualitative Research, Barcelona, Spain

We present the “evaluation planning process” of one of the National Strategies in Health, designed by the Spanish Government in the framework of the Quality Plan of the National Health System: the National Strategy of Rare Diseases.

Background
Spain is a State comprising 17 autonomies that hold exclusive competences in the field of health. Thus, strategies promoted at the State level suggest priority lines of action and recommendations that every autonomous community has to fit in its own circumstances and needs.

Objectives & Methods
To the extent that national strategies are the result of a thorough consensus among the scientific point of view and institutional interests of each territory, their evaluation planning work entails challenges and opportunities:

First it has been a methodological challenge. We have planned an evaluation focused on learning: as well as we try to account for the implementation of the objectives (performance indicators), we have also tried to surface questions related to the circumstances that have provided or hindered their achievements.

To do this, an on-line questionnaire has been administered among institutional representatives of each autonomous community within the rare diseases policy area. This questionnaire has been based on the graphical representation of the Strategy “theory of change”. On the one hand, completion of the questionnaire facilitated the discussion about appropriateness and relevance of the recommendation for action on each territory. On the other hand, has driven to an evaluation design sensitive to policies that involve polyhedric and complex processes of change.

Secondly, has been an opportunity in terms of institutional construction. Working from “theory-based evaluation” helped to clarify and critically relate the contents of the strategy and “what” is intended to achieve with it. Regarding to this, the evaluation planning process has revealed as an opportunity to permeate the strategy with directional effectiveness.

Keywords: Theory based evaluation; Learning based evaluation; Evaluation planning process; National Strategy evaluation

Capturing the impacts of leader and of measures related to quality of life in rural areas

European Evaluation Network for Rural Development, Helpdesk, Brussels, Belgium

The European Evaluation Network for Rural Development has prepared a Working Paper “Capturing the impacts of Leader and measures to improve the quality of life (QoL) in rural areas”. The paper provides methodological support for evaluators, managing authorities and others interested in assessment of impacts of EU-supported Rural Development Programmes (RDPs). There is a strong rationale for addressing these evaluation aspects together as the Leader approach to decentralised governance (integrated for the first time into RDPs 2007–2013) is a contributor to QoL in rural areas.

The main evaluation challenges for Leader and QoL include: assessing the “double scope” of Leader (both a process and generates products); the need to adequately define what is QoL in the RDP context; tackling the qualitative nature of the effects; identifying contributions from small-scale interventions; and noting the EU’s Common Monitoring and Evaluation Framework (CMEF) requires assessment of impacts ultimately at programme level. The working paper was developed by experts and is centred round four dimensions of QoL in rural areas – environment, socio-culture, economy and governance.

The hierarchy of indicators of the Handbook on the CMEF caters up to a certain point for external analysis of impacts. However, one of the strong points of the Leader approach is to involve beneficiaries and stakeholders in participatory evaluation, thus providing internal analysis. The Working Paper therefore suggests a three-step methodological approach including: (1) setting up a frame of reference including the monitoring information provided by CMEF common indicators as well as data from primary research (such as self evaluation and case studies); (2) the organisation of focus groups with stakeholders to collect complementary qualitative information; and (3) to compare and contrast the result with the baseline situation. Panels and evaluators would have to agree on a multi-criteria ranking format as a final outcome.

Keywords: Capturing impacts; Governance; Quality of life; Quantitative and qualitative methods; External and internal analysis

Current Evaluation Practice in the Czech Republic

J. Rem r

Presented paper tries to enlighten the current practice in operational programmes evaluation in the Czech Republic. The starting question is whether there are any specific patterns of how the specific elements of evaluation methodology are actually applied within the operation programme environment. Paper informs about the main results of the desk research that analyses the final reports from different types of evaluation studies (especially ex-ante, mid-term/ongoing and thematic evaluations). Conclusions of such research are supplemented and enriched with detailed knowledge build up on a series of in-depth interviews conducted with the authors of analysed evaluation reports (or generally evaluation research providers) and with the clients (i.e. managing authorities). Inclusion of both stakeholders creates excellent opportunity for finding the way how currently done evaluations could further be improved and how their utilisation for decision-making process could be enhanced.
Conducted research led to identification of five key topics that could be called as the typical characteristics of current practice. Above all it is (1) extraordinary emphasis that is put by both the providers and by clients on recommendations, (2) lack of feedback to already conducted evaluations, (3) rigor of conducted evaluations, (4) pseudo evaluative nature of required type of evaluations – especially in case of on-going evaluations and (5) low utilization of specific evaluative tools and techniques (i.e. logic models, meta-evaluation).

Research results could be found important mainly because they provide useful feedback relevant not only to evaluators in the Czech Republic but also in other EU Member States. Moreover, the methodology itself could serve as an inspiration for evaluators in other countries.

**Keywords:** Methodology of evaluation research; Operational programme; Meta-evaluation; Evaluation practice

---

**P-24**  
**Feminism and Gender in a Wider Global Political Debate Ugandan Perspective**  
**F. nk. Kiwalabye**  
*YCCWU, Social Work, Kampala, Uganda*

In Uganda and around the world, despite women’s hard work and progress, there continues to be a relatively rigid gender division of labor, between paid and unpaid work. Though women do participate in the political process in most countries, they are underrepresented in governments and their decision-making. Around the world, men dominate international security apparatuses and the making and fighting of wars. This paper addresses the real challenges to develop a full understanding of the nature and impact of gender mainstreaming in the modern world. Despite the importance of gender balance in global politics, gender is still not fully integrated in the global governance and academic study of international politics.

The main purpose of this paper is to demonstrate how developing innovative feminist methodologies is fundamentally affected by the unreflective interpretation of the concept of gender mainstreaming by governments and policy-makers.

The potential of mainstreaming lies in how information about gender issues is gathered in the process of mainstreaming gender. In practical terms, the potential of mainstreaming depends on how two key processes of translation are conducted. Firstly, how information about gender issues is gathered in the process of mainstreaming gender. Secondly, how gender-disaggregated data is then transformed into new terms, the potential of mainstreaming depends on how two key processes of translation are conducted. Firstly, how information about gender issues is gathered in the process of mainstreaming gender.

Secondly, how gender-disaggregated data is then transformed into new terms, the potential of mainstreaming depends on how two key processes of translation are conducted. Firstly, how information about gender issues is gathered in the process of mainstreaming gender. Secondly, how gender-disaggregated data is then transformed into new terms, the potential of mainstreaming depends on how two key processes of translation are conducted.

The main potential of mainstreaming lies in how information about gender issues is gathered in the process of mainstreaming gender. In practical terms, the potential of mainstreaming depends on how two key processes of translation are conducted. Firstly, how information about gender issues is gathered in the process of mainstreaming gender.

**Keywords:** Gender

---

**P-26**  
**Methodological challenges in evaluating a conference exhibit booth program at a federal health agency**  
**I. Zandberg**  
*Research Solutions International, Research, Rockville, USA*

This presentation will discuss methodological challenges in developing an evaluation for a centralized conference exhibit booth program at the National Cancer Institute. A study was conducted to assess feasibility of a full scale evaluation of this 20-year old program, which had never been evaluated before. Findings from this feasibility research indicated that no evaluation of a similar program had ever been conducted.

The program was both broad and deep, with diverse activities and audiences, including internal audience within the agency, and external stakeholders. The goals of the program were to: efficiently support the Institute in its communication and outreach activities; to increase participation of the Institute’s divisions, centers and offices in the program; to represent the Institute as accessible, trustworthy and credible source of cancer information; and to demonstrate the Institute’s commitment to and role in advancing science for the public. From this diversity of goals and audiences resulted varied outcomes that had to be taken into consideration in the program evaluation design.

It was concluded that the diverse activities of the program, and resulting range of expected outcomes and impacts called for a full scale multi-method evaluation that would go beyond the analysis of existing process data from Post-show Reports, Conference Summaries and Exhibit on Loan Tracking Data. A mixed method design involving both qualitative and quantitative assessments of program’s outcomes was proposed, including debrief staff survey questionnaire at the booth; debrief interviews with staff after conferences; ethnographic study including timing and traffic observation; evaluation forms from booth visitors, exit interviews with booth visitors. Proximal and shorter term measures were proposed to assess the impact of the program.

A logic model and evaluation templates that were developed for this project will be presented, and three research scenarios developed for three different levels of funding will be discussed.

**Keywords:** Evaluation design; Mixed methods design; Conference exhibit booth evaluation; Proximal measures; Logic model

---

**P-28**  
**The evaluation of training programs for capacity for the development. The challenge of integrating operational and strategic dimensions.**  
**M. L. Merino Cuesta**  
*La Agencia Estatal de Evaluación de las Políticas Públicas y la Calidad de los Servicios*

The evaluation of training programs for capacity for the development in the new paradigm. The challenge of integrating operational and strategic dimensions.

This article focuses on the evaluative approach to training programs for capacity for the development, under the requirements of the new paradigm. Mainly thesis is the need to integrate operational and strategic dimensions in the culture of the new management for results for the development and examines the key parts that require evaluation in the new model: integrated design, architecture and tools approved criteria aligned with international instruments. The study’s findings are based on a training program for capacity for the development of the Spanish Agency for Cooperation and Development Aid evaluated by the Spanish Agency for the Evaluation of Public Policy and Service Quality during 2009.
The Experience of Kirolbegi, Sports Observatory of Gipuzkoa, in the Study and Evaluation of the Regional Sporting Policies

J. Aldaz Arregui1, J. A. Arruza Gabilondo2
1 Kirolbegi. Sports Observatory of Gipuzkoa, University of The Basque Country & Grupo Iceberg Asesoramiento deportivo, San Sebastian, Spain
2 Kirolbegi. Sports Observatory of Gipuzkoa, University of The Basque Country, San Sebastian, Spain

In the present communication we show, in a succinct way, the evaluating experience of the guipuzcoan sporting system by “Kirolbegi, Sports Observatory of Gipuzkoa”.

In order to do this, we make a brief presentation of both Kirolbegi, Sports Observatory of Gipuzkoa and the regional context in Gipuzkoa, with the purpose of contextualizing the social-political framework in which it gets developed.

Next, we show Kirolbegi’s work studying its work philosophy and methodology in depth, clarifying the set of study dimensions that compose the guipuzcoan sporting system remarking especially one of the dimensions of the study related to Municipal Sports due to its special importance in the design of municipal sporting policies based of evidence, showing a SWOT diagram of it.

At last, we make a brief reflection that gathers some of the conclusions extracted from the thesis “The Practice of Sport and Physical Activity as a Liquid Habit of the Guipuzcoan Adult Population” by one of the authors of the present communication (Aldaz, 2009), trying to give response to if we might be facing Hyper (un)consciously arranged societies.

Keywords: Gipuzkoa – The Basque Country – Spain; Municipal sporting system; Observatory; Policy design

Exploring the impacts and outcomes of local regulatory services in UK local authorities

C. Celia, J. Tiessen
RAND Europe, Evaluation and Audit, Cambridge, United Kingdom

Background
Given the current economic climate, local authorities in the UK and elsewhere are finding that they are increasingly expected to demonstrate the benefits and added value generated by their work and their local regulatory services (i.e. services such as fire safety, food standards, trading standards and environmental health) in particular given the impacts these can have on both businesses and consumers.

Objectives
The aim of this work was to contribute to developing a robust evidence base to demonstrate the added value and benefits of these services. The methods proposed had to be easily applicable to all local regulatory services and be practical enough to be of direct relevance and use to local authorities.

Methods
This work, based on a theory of change approach, developed a toolkit for local authorities to help them develop their own contribution stories. Practically, we used an adapted version of a logic model, a “causal pathway”, as its starting point along with a dashboard methodology to present indicators for each stage of the pathway as shown below:

The data-gathering methodologies underpinning the development of the pathway and dashboard included: scoping interviews with local authorities, an online survey of all local authorities in England and Wales and a series of workshops with local authority staff to work through causal pathways and dashboards in different areas of local regulation in order to check their practicality and applicability to local authorities “on the ground”.

The suggested poster will present the methodological approach chosen, and provide practical examples of how the tool was used by local authority regulatory services in the UK.

Keywords: Impact; Outcome; Local regulatory services
**P-32 The Challenges of M & E in the Domestic Taxes Division of Zambia Revenue Authority**

*I. Njovu*
Zambia Revenue Authority, Domestic Taxes, Lusaka, Zambia

**Background**
The Zambian government is implementing the Paris Declaration on Aid Effectiveness. Much has been done during the past few years in getting the government to implement the principles of ownership, managing for results and mutual accountability on the national level. This is mainly at the ministerial level. However, much also need to be done on the lower level of government in local government, government departments and agencies.

**Objectives**
The aim is to show the challenges of managing for results and accountability at a macroeconomic level, especially in revenue or tax policy implementation. There is a new directorate of Designing and Monitoring in the Zambia Revenue Authority. My presentation will be on the challenges that arise out of the change management process and the implementation of a monitoring and evaluation system at a lower level of government. This is especially when an external implementing consultant is only available for a very short time.

**Methods**
The method to be used will be mainly a case study method.

**Conclusion**
I am hoping that other monitoring and evaluation practitioners will learn from our experience. I also hope that other departments of government implementing the Paris Declaration will also learn a lesson or two from our experience and need not re-invent the wheel. There may be others who also can share with us their experiences and help us resolve some of the challenges we have faced.

**Keywords:** Zambia Revenue Authority

---

**P-33 Learnings from Evaluation of IDPs project**

*F. Khan*
Catholic Relief Services, Program, Islamabad Pakistan, Pakistan

**Introduction:**
Catholic Relief Services – Pakistan Program is working in Pakistan since 1954 and contributed to almost every Emergency that struck the country e.g. Earthquake in 2005, floods in 2007, Earthquake Balochistan 2008 and IDPs crises in NWFP etc.

CRS is a learning organization and now CRS Pakistan always welcomes new ways for improvement for making its Monitoring and Evaluation country team stronger and disseminates the information and good practices to its local partners. This upcoming conference will help me and my country office to move forward with new ideas and suggestion to build a better M&E team on the country level. There is significant lack of M&E people and it is hard to retain the less number of specialized M&E people in our country offices, due to better opportunities in other organizations. I believe this conference will enable me to help my country office and other staff in developing new ways to improve our M&E strategies and equip our selves with more improved and tested ways used in other regions of the world.

**Presentation:**
I will also share experiences from my country, in the shape of presentation, with diverse group of workshop participants. These experiences will include the learning’s we have done so far and that how it helped our organization in moving forward with a better and improved practices. The presentation will present learning from the recently conducted Evaluation by CRS Pakistan of its IDPs crises emergency response program, which was one of the most complex emergencies ever struck this region.

I believe it’s a better opportunity for me to share CRS country program M&E experiences and learnt from the conference group participant’s experiences and their views for a better organizational future strategy. Last but not the least, accountability and efficiency of our work towards organizations and beneficiaries.

**Keywords:** Learnings; Evaluation; IDPs; Project; Sharing

---

**P-35 How to measure, monitor and evaluate progress and impact of climate change related interventions in a development context?**

*C. Roehrer*
DfID – UK Department for International Development, Evaluation Department, Glasgow, United Kingdom

Given the political importance of climate change internationally and in the UK with increasingly larger sums being invested (e.g. $30 bn for Fast Start agreed in Copenhagen for 2010 and 2012, $100 bn annually by 2020 – and the UK having committed to spend £1.5 bn on Fast Start alone, clarity, transparency, oversight and evaluation are becoming ever more important. But challenges regarding evaluation of climate change are still huge: There is a lack of knowledge on how to best measure accomplishments and outcomes in this complex field of work. In order to cope with climate change, non-linear and systemic changes are needed (hard to measure) and we have to deal with “slow onset challenges”, irreversible and unstoppable changes in the climate (e.g. sea level rise, drying trends, salinization) alongside climate variability. Beyond that we still lack agreed definitions of for example what makes adaptation different from "just good development" and we basically do not yet have an agreed methodology for measuring adaptation outcomes.

Evaluation challenges on climate change in a development context will require methods that address risk and uncertainty, methods that

– rather consider complex system effects (backward and forward linkages) than linear relationships between means and goals,
– take explicit account of unintended consequences and side effects,
– include economic, social, distributional, environmental and security indicators.
Evaluation into practice: from lessons learnt to lessons used.

A. Salem¹, D. Innocenti²

¹ Danish Refugee Council, Horn of Africa and Yemen, Nairobi, Kenya
² United Nations International Strategy for Disaster Reduction, Europe Regional Office, Geneva, Switzerland

The Project Cycle Management and the Results-Based Management approach identify as critical the identification of a programme or a project, the moment in which objectives are set, outputs agreed and activities and (consequently) indicators of output and outcome shall be identified.

In this sense it is crucial to build a Management Information System (MIS), which includes the Monitoring and Evaluation (M&E) framework, so that procedures to guide the recommendations formulated by the previous evaluations are mainstreamed in the identification and subsequent implementation.

Indicators give to the Management the “yardstick” to measure results but its the development of a functional MIS which allows Managers to effectively use their “yardsticks” quantifying what has been achieved against resources used.

This presentation deals with the latest concept: building a conducive MIS and monitoring and evaluation framework to assure that lessons from evaluation are used by managers.

Bringing into the discussion “first hand” experiences, we are documenting what the MIS of a community based project implemented in 2008–2009 in Somalia has developed in order to allow its Management to prepare the ground for the evaluation which can appreciate the impact achieved making recommendations for future interventions which are effectively used in implementing the project activities.

Establishing determinants of behavior relating to improved governance among constituents in Nigeria

M. Cohen

RTI International, M&E, Research Triangle Park NC, USA

Multiple theories propose and test factors correlated with health behaviors and thus improved health, and proven determinants are used in international development to shape program theory and guide implementation. One recognized determinant of improved health in developing and transitional countries is good governance. However, we are at the beginning of the search for understanding the factors that may predict attitudinal and behavior changes required for instituting and sustaining good governance.

The need for projects to measure and use data for evidence-based decision making requires that evaluators be able to provide recommendations based on verifiable data obtained through established techniques. In many projects, the role of the evaluator is limited to measuring impact after activities are completed. While this is indeed an important role, it often means that some of the most important programmatic decisions at, or near, startup are not based on evidence, but assumptions. Perhaps even more than in the context of a health project, the determinants of a desired behavior for good governance may change from one country to the next, from one region to another.

The inputs and messages required for constituents in Cambodia will be different for those in Colombia.

In preparation for a project aimed at strengthening the foundations of democratic governance in targeted states of Nigeria, a framework has been developed to help provide initial guidance during startup, combined with an approach for gauging impact at its conclusion. The methodology combines quantitative assessments of access to critical inputs with psychological and social motivations for engaging in desired behaviors. By correlating specific potential determinants – such as outcome expectations and social support – with the performance of key behaviors, we can learn which attributes lead to this behavior, and can target the program’s inputs and resources accordingly.

Disengaging Evaluation from Monitoring: The Challenge of Balancing Accountability with Learning in Development Projects in Afghanistan

L. Bulosan

The Asia Foundation

Europe contributed almost 48% of the total development aid received by Afghanistan from 2002–2008 and yet the country still ranks 174th in the Human Development Index. Almost all the unilateral and multi-lateral donor groups and development contractors, for profit and not for profit international development organizations converge in Afghanistan. And yet most Afghans, especially women and children, still endure conditions of hardship and extreme poverty. Given the nexus between development and security, the effectiveness of aid has a major impact on peace and stability in Afghanistan. However, a survey of the monitoring and evaluation systems of development projects in the country indicate that there is more emphasis on the accountability (monitoring) and less on learning (evaluation) that would deepen our understanding of what worked, what did not, why and how. Often times, because M&E are said on the same breath, there is failure to recognize that evaluation
is the more important variable in the equation. Projects focus is on getting the logical frameworks right and and collecting data for the corresponding performance indicators while the evaluation research design is almost completely neglected until it is time for mid-term or end-term evaluation. Thus, there is a need to disengage evaluation from monitoring in the mindset of development practitioners. Monitoring tells us how busy we are but not how effective we are. Evaluation does. In a highly volatile environment such as Afghanistan, the urgency to emphasize evaluation and develop just-in-time cost-efficient evaluation tools that are continuous and not periodic is imperative.

In Afghanistan, when will we ever learn of the effectiveness of aid in mitigating conflict? Not until development projects give equal attention to accountability and learning. Not until we advocate for funding agencies to require a separate monitoring plan from an evaluation plan. Not until we disengage evaluation from monitoring.

**Keywords:** Afghanistan; Stability; Evaluation; Aid effectiveness; Developing

---

**P-41 Direct Budget Support: Potential and Limits of Evaluation of Highly Complex Development Aid Interventions**

*A. Larcher*  
*ISCTE*

After decades of development cooperation, the question of whether aid works is ever more pressing. Some call for aid to be stopped altogether, while the international donor community, as a response to critics, has put forward a reform agenda to increase aid effectiveness. This agenda has led to a shift in aid practices. It confers a more central role for evaluation, seen as a key to analysing the impact of aid and it has led to new aid delivery modalities such as direct budget support, conceived to help resolve many of the problems of aid inefficiency.

This approach is much more complex, than the previous project and programme approaches, and presents several challenges to evaluation. This paper will ask, taking budget support as a case study, whether current evaluation practices can help us understand the complexity of the impacts of aid and where they are failing. What are the impacts which are invisible to evaluators given current evaluation practices and which determine the relevance of development interventions? Which are the unintended consequences of external interventions? How can evaluation practice be transformed so that it is useful in capturing (1) the dynamics of societies they are seeking to transform and (2) the interests of all different stakeholders in aid industry and how they affect aid impacts and (3) the cumulative impact of all aid interventions?

Evaluation is not adequately practiced has the danger of leading to interventions being misunderstood and of not providing us with the necessary information for change.

**Keywords:** Development aid; Direct Budget Support

---

**P-42 Monitoring and Evaluation Dynamics of Capacity Building in Developing Countries: Measurement Challenges**

*J. Pradhan*  
*National Institute of Technology, Humanities and Social Sciences*

Over many decades billions of dollars have been invested in conducting projects, providing technical assistance to the countries where the project was developed, and educating and training high-level specialists in developing countries; yet lack of capacities to design, plan and implement policies and reforms remains today arguably one of the most serious obstacles to the implementation of health and education projects in developing countries. Many donors and sub-donors clearly want to be able to show how ‘X’ amount of resources invested resulted in “Y” amount of improvement, change or development. The problem lies in the “as they intended”. Donors and sub donors tend to create a tremendous pressure for those whom they support to plan and measure much more than they could possibly, intelligently predict. The consequences of this pressure are manifested through huge efforts by implementing actors to show results in order to comply with sanctioned programming mandates and to demonstrate a direct relationship between cause and effect. This situation lays the foundation for much of the linear thinking that is endemic in development programming. With this background, this paper aims at to discuss four monitoring and evaluation (M & E) dimensions of capacity building viz; a) Static development versus development in motion; b) Linear versus complex adaptive systems (CAS) thinking, programming and measurement; c) Attribution; d) Donor accounting focus versus open learning approaches

---

**P-43 Evaluation of the effectiveness of the United Nations Framework Convention on Climate Change Secretariat in capacity-building in developing countries**

*S. Kozlov, K. Kozlov*  
*Forest Stewardship Council (Russia), Novosibirsk branch, Novosibirsk, Russia*

**Background**

Capacity-building is considered to be an essential mechanism in assisting developing countries to address climate change and its effects. Under the United Nations Framework Convention on Climate Change, the Subsidiary Body for Implementation is charged with advising about capacity-building in developing countries (Article 9), while the Kyoto Protocol commits Parties to the Convention to cooperating in the strengthening of national capacity-building. Secretariat assists the Parties to implement the principles of the Convention through the development and implementation of international, national and regional strategies, policies and programmes based on SBI recommendations and Conference of the Parties’ decisions.

**Objectives**

This evaluation design of the Capacity-Building framework aims to assess the progress made by the Convention Secretariat in mitigating climate change and adapting the developing countries to its effects during the period 1995–2009. The evaluation represents a results-based programming approach to assess functions of the Secretariat. The evaluation is designed to make a statement whether the Framework for capacity-building in developing countries is effective in assisting developing countries in the implementation of the Convention and/or the Kyoto Protocol.
Hybrid innovation: Monitoring and Evaluation in corporate social responsibility development initiatives

C. Elkins
RTI International & Duke University, Sanford School International Development Policy, Hillsborough NC, USA

Background
In contemporary international development, businesses act in conjunction with traditional development actors, and alone. Corporate social responsibility (CSR) design and management practices, however, often overlook decades of learning achieved through traditional public sector and internationally-led development projects. Typical reviews of CSR projects focus on return on investment to the firm, its stakeholders, or the so-called triple bottom line (people, planet, and profit). Assessments do not rigorously evaluate or measure the nature and extent of CSR interventions’ effects on targeted communities or broader populations. When attempted, evaluation may inadvertently skirt social science research standards with respect to informed consent, privacy, or other ethical concerns.

Opportunity
Where CSR initiatives intend to build social and economic capacity, promote engagement in community concerns, or explicitly strive toward sustainable local development, they offer alternatives that can catalyze social and economic progress (development). CSR actors can operate as donors, partners, implementers, and/or links in the supply chain. CSR actors can adjust scopes, partners, and funding more flexibly than typical development bureaucracies. At the same time, CSR explicitly works from a profit-minded perspective rather than altruism, so dissemination of development impact will often be received skeptically or considered publicity, not evidence.

Challenge
With their unique position and flexible contributions, private sector actors can contribute not only to development, but to development theory and pragmatic knowledge. A hybrid model of monitoring and evaluation will help capitalize on shared stakeholder interests and development goals, motivating more constructive corporate engagement and critical resources through tough economic times.

Application to Practice
Attention to international development norms and M&E standards strengthens the credibility of CSR initiatives in traditional development circles and contributes directly to the impact, effectiveness, and sustainability of interventions. This paper presents a practical model of M&E/evaluation implementation adapted specifically for use in hybrid CSR development interventions.

Keywords: Monitoring and evaluation; International development; Private sector; Impact
P-47 Building a culture of performance and evaluation in Punjab, Pakistan: A Case Study and Lessons Learnt

S. Ali
Earthquake Reconstruction and Rehabilitation Authority, Financial Management Wing, Islamabad Pakistan, Pakistan

Punjab is the largest province in the Federation of Pakistan and the largest sub-national economy contributing about 60 % of the national Gross Domestic Product (GDP). Punjab's economy in the 1990s was not expanding sufficiently for creating enough jobs for a significant reduction in the incidence of poverty.

A new economic policy framework was put in place in the province of Punjab in 2004 the main thrusts of which were embodied in Punjab’s Vision 2020 which broadly spelled out a host of interventions in economic, social and governance realms. A strong reform agenda built around: (i) improving governance; (ii) reforming the fiscal and financial management system; (iii) creating a more supportive environment for private sector-led growth; and (iv) improving the delivery of public services was put in place.

The new policy direction required a strong focus on monitoring and evaluation calling for development of suitable databases and strengthening of economic research and statistical institutions. These were important components of the Value Chain leading to good economic policy making in the province.

To this end, inter alia, an annual economic survey report was instituted as part of the efforts to build internal capacity for effective monitoring and evaluation of the development process in the province. It was hoped that this annual process would strengthen the feedback linkages with the line departments, identify data limitations and provide a forum for internal debate and discussion. A large multi-topic household survey using the MICS methodology was undertaken with over 90000 households as a basic instrument for introducing an evidence based policy process in the social sectors. This paper/presentation would be a case study of this experience documenting successes, failures and lessons learnt from these endeavours.

P-48 Livelihood Challenges in Development – an Evaluation of Involuntary Displacement and Resettlement Project in Central India

U. Chatterjee Saha
Xavier Institute of Development Action and Studies, Department of Health Care, Jabalpur, India

Most large-scale development projects in India have been planned and executed through a centralised, hierarchical and capital-intensive approach, which has assumed that technology has the answer to all development problems. Centralised command and control system has perhaps been successful for achieving project objectives in terms of physical commodities like generation of thermal power, hydro-electric power, irrigation canals. However, the approach has almost completely failed in resettling and rehabilitating the displaced population who have suffered intolerable injustice and trauma.

The Bargi Dam was built across Narmada, the fifth largest river of peninsular India. In the process of construction about 162 villages were affected out of which 82 villages were completely submerged. This study tries to understand the success of R&R programme in livelihood restoration among 693 Project Affected Families (PAFs) from 22 villages of three districts in the state of Madhya Pradesh.

The main objective of the study is to assess the extent of impoverishment risks of the PAFs due to construction of Dam. It also evaluates strength and weakness of the R & R policy and implementation strategies of India Government.

The bivariate and multivariate analysis illustrates that low caste people are most vulnerable to the process of displacement as they are illiterate, marginalised and only skilled in traditional farming. The displacement led to massive land loss and economic transition from agriculture to non-farm labours, business etc. where they have least expertise. Although 90 percent of the respondents could procure the compensation but most of the money was used in mere survival until they fully adapt with new life. Thus advocacy and training regarding self employment, promotion of children’s education, small scale and multiple farming, water resource management and micro-level saving may help them to come out of their dilapidated condition and develop a sustainable livelihood system.

Keywords: Project Affected Families; Displacement

P-49 Efficiency and sustainability of development initiatives through coordinated evaluation mechanisms

D. Ter-Stepanyan¹, A. Avagyan²
¹ Civic Development and Partnership Foundation, Monitoring and Evaluation, Yerevan, Armenia
² Counterpart International Rep in Armenia, Civil Society, Yerevan, Armenia

A number of development programs are implemented in Armenia by national and international organizations. Between 2001 and 2005, Armenia received nearly 1,2 E billion of net official development assistance flows, about 235 million yearly. Many of the implemented programs are based on a through need assessment. However, not all of these programs are implemented in addressed, targeted, and efficient manner and are mostly implemented uncoordinatedly and in a scattered manner. Moreover, hardly any program conducted in past has been evaluated as regards to its long-term results, sustainability and impact for the community and the region.

As show the investigations sustainability of development programs is becoming into the focus of attention for development agencies. In order to ensure sustainability, it is important that development initiatives involve need assessment and follow-up impact assessment components. According to discussion with local organizations, the weakest part of project implementation is achievement of the intended impact and the sustainability, which could have several reasons. Preliminary findings show several reasons for that: (1) many donor agencies are coming with their own agenda and goals that very often are not adapted to the local necessities and situation, (2) sometimes the project implementation process lack transparency, accountability and openness for community participation which brings to less effectiveness and sustainability of implemented projects, (3) after project closure rarely a follow-up monitoring and evaluation is implemented to insure that project results are at place.
Using Performance Indicators to Promote Informed Public Discourse on Educational Achievement in the Anglophone Caribbean

M. Taylor
University at Albany, Education Administration and Policy Studies, Albany New York, USA

Since gaining independence, many of the former colonies of the UK have expanded their local education systems in the hope that doing so would promote democratic governance and national development. At the Jomtien and Dakar Education for All summits, the global community intensified its call for countries, particularly developing nations, to provide quality educational opportunities for their denizens. Governments in the Anglophone Caribbean have responded to this call. They have established the goal of providing, in the near term, universal access to Secondary Education (USE) of high quality.

Despite the apparent commitment by many of these governments to their stated goal, there are some who are doubtful that the various education systems have the capacity to make effective progress in this regard. Debates are frequent, but are typically not grounded in evidence. The indicators currently generated are not adequate to provide a full status on the impact of secondary schooling and does not allow for the level of monitoring and evaluation needed to assess the effectiveness of the efforts being made to achieve the stated goal of providing quality USE. As a result, this paper proposes that countries of this region adopt a new set of four indicators which incorporates the following; two variables which focus on the level of student participation in education, one which measures the quality of student performance in those examinations and the final indicator which amalgamates the previous three indicators to develop an index which reflects the effectiveness of schools in providing USE of quality. Having an objective assessment of the status of education and being guided by the indicators to ask relevant questions about the effectiveness of schools, districts and countries will encourage various publics to engage in informed discourse, focused on how improvements can be made to realize greater achievement.

Keywords: Anglophone Caribbean, Universal Secondary Education, Exit examinations, Informed discourse

Impact evaluation of small grants allocation in environmental preservation in Central Africa: Case study of Lake Chad Basin Commission

S. E. Yakou Djiam
Gitech Consulting

The present paper derive from the results of a monitoring and evaluation mission of 25 micro projects financed by the International Union for Nature (IUCN) in Cameroon, Niger, Nigeria and Chad conducted from July to September 2007. These concerned communities’ micro projects of the “Diagnostic of the boundary and the North part of Lake Chad” project, an activity of the Project called “Reversal of Land and Water resources degradation trends in the Lake Chad Ecosystem” executed by IUCN. The main objective of the mission was to verify the level of achievement of activities with the aim of identifying the critical constraints and adequate solutions to be used in the second phase of the project. Several lessons were carried out from these dynamics: (i) small grants subvention to local projects is a key issue for the sustainable management of natural resources and for environmental preservation; (ii) the earlier knowledge of a community is an asset for the success of a project, this means that their development problems are known before; (iii) the employment of pedagogic tools and others illustrated tools (with pictures) is a strength to facilitating the adoption of associated techniques of innovations for the development in the case of illiterate target; (iv) in tree plantation, the choice of species to be used is related to their contribution to secondary needs of communities (multiple use’ tree for example); (v) setting up of project’ activities in rural milieu when taken into consideration the allocated time by populations to their agricultural activities is a guarantee which could guide the success of the project; (vi) the implication of local authorities and others influenced members of beneficiaries’ communities, in an executed development project as strategic communication and sensitization agents move to the stimulation of effective participation of the entire community.

Keywords: Diagnostic; Reversal trends; Degradation

Values in Evaluation: reflections over a case study

P Rodríguez-Bilella1, R. Monterde Díaz2
1 Universidad Politécnica de Valencia, Dpto. Proyectos de Ingeniería, Valencia, Spain
2 Facultad de Ciencias Universidad de San Juan, San Juan, Argentina

Although the evaluation of public policies is subject of a growing interest in Latin America, there are problems with the design and implementation of evaluations, as well as with the limited use of their results. In many cases, the evaluations have more to do with the generation of descriptions, and less with the valuation of these activities and its use in order to improve planning and decision making. These points are explored from a case study of the evaluation of a rural development program in Argentina, emphasizing the process of negotiation and consensus building between the evaluators and the official in charge of approving the evaluation report. The lessons learnt from the experience point to the generation and consolidation of a culture of evaluation in the region.

Keywords: Culture of evaluation; Values; Negotiation
P-54  Evaluation of an urban regeneration policy: how to improve social cohesion in neighborhood?

M. Simo Solsona
University of Barcelona

The main objective of this poster is to present an evaluation of regeneration policy implemented in Catalonia (Spain), concretely the Llei de Barris (Neighborhood Policy). The main purpose of this policy is to promote an integral regeneration process in these areas to avoid urban deterioration and improve the conditions of life of residents. In order to do that, the policy is composed by social, economic, urban and environmental interventions. Concretely, actions are focused to restrain demographic problems (such as high concentration of ethnic population, ageing), economic and social problems, for example stigmatization and expectations to move away of the neighborhood.

The evaluation is an opportunity to analyze why some neighborhoods are more predisposed to suffer a deterioration process and which are the main origins of this process. Moreover, the evaluation could be a good instrument to test how well-being of the residents is after these interventions program, if they improve their basic needs and their future expectations, as well the attachment to the neighborhood.

The evaluation methodology applied has a global and integral character due to complexity of the policy such as wide range of actors, multiple actions executed, diversity of neighborhood and different evaluation culture from professionals. The main contribution of this poster is the construction an indicators system to evaluate the different type of actions (related to different objectives): contextual indicators, monitoring indicators (sufficiency, relevance and pertinence) and finally, effects indicators (realizations, results and impacts, correlated to effectiveness, efficacy and efficiency).

Summarizing, the poster analyzes up to what extent the policy has influenced the social cohesion in these neighborhoods, the quality of life of residents and how these neighborhoods are more embedded in the city, slowing down social and spatial segregation of these places.

Keywords: Policy urban regeneration; Indicators system; Integral evaluation

P-55  Challenging champions: Evaluating good governance in motion

C. Elkins
RTI International, Director M&E/Eval, Hillsborough NC, USA

Background
Good governance requires more than compliant institutions in a participatory framework. Roots of and sustenance for governing with consent and participation of the governed rest at least equally with civil society. Transparent and accountable government is truly defined in the active participation of multiple actors within the relevant institutional framework. Yet to what extent can civil society effectively entrench standards of good governance without significant opening, clearly meaningful encouragement, or minimally grudging participation by those who govern?

Opportunity
Opportunities for grounding good governance in Guinea (West Africa) may appear limited from the center, but have significant potential to thrive at the grassroots level. Initiatives underway prior to disruptions of 2009 have taken new energy from recent events and decisions, including elections in a series planned to start with a Presidential contest 27 June (legislative and local elections to follow).

Challenge
With a limited budget and 18-month timeframe, a USAID-funded project (Faisons Ensemble) will undertake facilitation of demand-driven community action strengthening effective governance and results in locally-prioritized challenges in health, education, and agriculture. Instead of a grants focus, Faisons Ensemble requires initiative from demonstrated achievers working at the base. Initial program support facilitates community self-identification of their own priorities and self-selection by the most motivated and effective local volunteers to mobilize needed group action.

Application to Practice
Attention to champions and personal leadership has a rich tradition in political science, but far less so in political economy and international development theory. Yet contemporary development practice often leans heavily on expectations that local motivation and leadership are essential factors in successful achievement of results, and hope for sustainable impact. This paper presents a symbiotic approach to project and M&E/evaluation design that can begin to test these hypotheses through conscious capture of relevant evidence throughout the intervention process.

Keywords: International development; Governance; Africa; Evaluation; Monitoring and evaluation
List of Speakers

A
Abma, T. 58, 60
Abreu Nogueira, J. 72, 73, 98
Abreu, D. 190
Afonso, P. 221
Aigner, D. 95
Albuquerque Bezerra, L. C. 29, 216
Aldaz Arregui, J. 223
Ali, S. 227, 228
Almeida Alves, C. K. 29, 216
Alvarez, P. 94, 201
Andrei, T. 160
Andreozzi, A. 102
Antunes, P. 147
Arcand, J. L. 65
Aredo, D. 197
Arenes, Y. H. 199
Arruza Gabilondo, J. A. 223
Astbury, R. 104
Astedt, E. 102
Atanesyan, K. 177
Avagyan, A. 228
Aw-Hassan, A. 198
Azzam, T. 46, 121

B
Bal-Doebel, L. 42, 192
Balenzano, C. 35
Barbulescu, I. 210
Barrett-Smith, J. 70
Baskakova, J. 183
Baur, V. 58
Bayley, J. S. 83
Bayley, S. 150
Beaufoy, G. 142
Beaulieu, R. 155
Bednarikova, Z. 45
Belanger, D. 188
Belbase, K. 82
Beney, T. 147
Berden, K. 106
Beres, A. 10
Berg, R. 116
Bergschmidt, A. 6
Bernard, T. 65
Boehme, K. 64
Bohm, S. 72
Bolli, M. 9, 45
Bonifazi, A. 63
Booolsen, M. W. 129
Borek, A. 220
Borenstein, J. 190
Bossuyt, J. 17
Boyle, R. 172
Brans, M. 103
Brejc, M. 61, 96
Brisebois, I. 80

Broegaard, E. 135
Brouselle, A. 13, 33
Brown, D. 179
Bruhin, E. 114
Bruni, G. 202
Brusset, E. 134
Buffoni, B. 18
Bulosan, L. 225
Burakiene, D. 160
Bush, K. 69
Bustelo, M. 26, 94, 201
Byeon, S. C. 163

C
Cabria, A. 93
Caetano, A. 42, 85
Cagliero, R. 217
Calabria, A. 131
Callerstig, A. 201
Canu, E. 124
Caputo, E. 35
Carden, F. 30, 121, 146
Cardona, A. 132, 220, 221
Cardoso, G. 190
Carlhed, C. 27
Carter, P. 36
Carvalho, A. 42, 192
Casavola, P. 89, 158
Cavaco, N. 85
Cedzynska, M. 185
Celia, C. 223
Cervantes, M. 133
Chaplowe, S. 131
Chatterjee Saha, U. 101, 144
Cheng, A. L. 25
Chomel, J. 126
Chrispin, P. P. 190
Ciampi, S. 9
Clemenhenagen, C. 179
Clercxx, E. 178
Cloete, F. 195
Coden, N. 132, 220, 221
Cohen, M. 151, 179, 225
Cola, E. 190
Congiu, D. P. 122
Conner, R. 121, 198
Contandriopoulos, D. 33
Conzimu, L. 11
Cordero, C. 181
Corinial, F. 128
Costa, A. S. 202
Costa, T. 85
Cousins, J. B. 32
Crespo, R. 93, 132, 220, 221
Cristiano, S. 45, 46
Crivelaro, R. 39, 218

Cruz, M. 49, 92, 190
Cruz, M. M. 128
Csengodi, S. 10
Cugelman, B. 186
Cummings, A. 80
Cummings, H. 80, 109, 136
Cunningham, F. 59

D
Da Silva, S. 146
Dahler-Larsen, P. 26, 53, 74
Davidsdottir, S. 96, 97
Davies, I. 30
Davies, R. 110
De Groot, D. 107
De Lancer Julnes, P. 219
De Peuter, B. 103
De Ruiter De Wildt, M. 197
De Schepper, G. 205
De Vet, J. M. 106, 167
Deblois, M. J. 80
Decottelli, P. 190
Decottelli, P. V. 92
Delarue, J. 65
Denison, E. 116
Denvall, V. 90
Dessers, R. 44
Devischer, S. 203
Diaz, J. M. 164
Dietz, T. 107
Diop, N. 179
Doerflinger, A. 215
Dolce, A. 119
Donaldson, S. 41
Dos Santos, E. M. 128
Duggan, C. 30, 69
Duchene, V. 76
Dunnett, G. 137

E
Eberhard, R. 8
Edelmann, S. 134
Ehler, C. 203
Ehneström, C. 202
Elkins, C. 151, 227, 230
Elliott, C. 32, 162
Ellis, G. 105, 137, 175
Ellis, R. 140
Elsworth, G. 164
Erdebrir, T. 184
Eriksson, B. G. 157
Eronen, A. 144
Errejon, J. A. 164
Espinosa, J. 93, 201
Esteban, V. S. 166
Etta, F. 83, 201, 219
Etting, D. 108

WWW.EUROPEANEVALUATION.ORG 231
Rieper, O. 78
Rigout, F. 91
Riisom, T. 72
Rijneveld, W. 107
Rischke, R. 195
Rissotto, A. 131, 190
Rist, R. 32
Rodriguez-Ariza, C. 181
Rodriguez-Bilello, P. 229
Roehrer, C. 224
Rooijmans, M. 28
Romeo, B. 57
Romano, B. 216
Romero Martinez, M. 153, 154
Rosselli, A. 158
Rugh, J. 48, 83
Ruhland, S. 24, 216
Rune Jensen, M. 18
Rutten, R. 110
Rydevik, I. 54
Saar, Z. 14
Salem, A. 225
Samek Lodovici, M. 158
Samaco, I. 29, 216
Sanders, D. 203
Sansonetti, S. 158
Santos, E. 49
Santos, S. B. 92
Sanz Luque, B. 5, 159, 201
Sanz, B. 201
Saraswat, M. 5
Sarid, A. 14
Sarmento, E. 196
Sartori, D. 148
Sarts, M. 111
Savaya, R. 14
Scardigno, A. F. 189
Seabi, M. 176
Segone, M. 5, 171
Severati, P. 138
Shah, N. 14
Shim, J. M. 218
Shina, I. 14
Schaffner, S. 203
Schanne, M. 38
Schaumburg-Müller, H. 124
Schiefer, U. 42, 192, 216
Schiller, S. 142
Schipper, K. 60
Schmidt, W. 40, 134
Schnaut, G. 142
Scholze, T. 175
Schroll, M. 85
Schuh, B. 142
Schüle, R. 199
Sieb_Pos, M. 130, 195
Silvani, A. 40
Simerska, L. 51
Simó Solsona, M. 230
Simons, H. 43, 207
Sindzingre, A. N. 34
Singh, N. 108
Siri, A. 49
Sirilli, G. 40
Slee, W. 221
Snitstveit, B. 116
Sohnman, A. 218
Spampinato, M. 80
Spapo, A. 77, 105
Speer, S. 104, 150, 172
Spera, M. 190
Spiessens, S. 205
Spurk, C. 38
Sridharan, S. 146
Stame, N. 136, 150, 172
Starling, S. 186
Stebenkova, L. 183
Steurs, G. 76
Stewart, D. 121
Stockmann, R. 16, 26, 136
Streicher, J. 162
Sudarshan, R. 173
Suelzer, R. 31
Sullivan, F. 153
Suszek, K. 185
Svensson, L. 202
Svoboda, D. 15
Sylvest, J. 63
ta Larsson, B. 70
Tagle, L. 9, 89
Tall, O. B. 171
Tamminen, S. 106
Tamondong Giraud, S. 217
Tanese, A. 141
Taranghio, F. 45
Taylor, B. 8
Taylor, M. 229
Teixeira, P. 47, 147
Téllez Rojo, M. M. 133, 153, 154
Temminck, C. 127
Terberger, E. 195
Terraseca, M. 202
Ter-Stepanyan, D. 228
Terzi, A. 141
Thieba, D. 71
Thomas, L. 153
Tiessen, J. 223
Ton, G. 197
Toropova, N. 184
Torres Pereda, M. P. 133, 153, 154
Torres, R. 92
Torrigiani, C. 122, 211
Toudal, M. 137
Toulemonde, J. 26, 63, 75, 138, 204
Tourmen, C. 75
Tugny, A. 118
Tunali, N. 184
Turgeon, J. 67
Tuzi, F. 40
Übelhor, M. 125
Udsen, L. 99
Unceta, K. 181
Usukiyal, P. 144

V
Vaessen, J. 116
Vahlhaus, M. 33
Valmarana, C. 35
Van den Haute, B. 205
Van Hoorebeeck, B. 205
Van Laere, F. 205
Van Ongevalle, J. 127
Van Overbeke, M. 119
Van Stolk, C. 131
Vandeputte, L. 205
Vandersmissen, M. 205
Vedung, E. 115
Ventura, M. 92
Vergani, A. 101
Verlet, D. 205
Vermoesen, E. 205
Vestman, K. O. 157
Vilarins, L. 39, 218
Virtanen, P. 100
Visse, M. 60
Voigt, H. 195
Voskes, Y. 58
Vougioukalou, S. 67
Voylokov, T. 183
Vrangbæk, K. 117

W
Waddington, H. 116, 198
Wagner, N. 65
Walczak, B. 215
Walker, D. 208
Wallin, S. 86
Walsh, P. 200
Weddwood, H. 171
Weinspach, U. 221
Wennberg, M. 148
Wessels, J. 8
Westhorp, G. 104
Westin, L. 27
Wikner, M. 54, 87
Wille, S. 178
Williams, B. 147
Williams, K. 52
Wimbush, E. 13, 204
Winckler Andersso, O. 135
Y
Yakeu Djiam, S. E. 229
Yong Protzel, I. 212
Yoo, S. J. 25
Z
Zaal, F. 107
Zacarreri, M. 82
Zandberg, I. 222
Zanetti, B. 45
Zatoms, W. 185
Zaveri, S. 173
Zeyen, P. 114
Zhai, Q. 29
List of Keywords

A
Absorption ....... 170
Accompanying Evaluation ....... 199, 216
Accompanying research ....... 8
Accountability ....... 34, 89, 103, 112, 187, 206
Accountability of resources ....... 71
Accountability to donors ....... 206
Action-research ....... 127
Active labor market policy ....... 203
Activism ....... 147
Actor types ....... 107
Actors-perspective ....... 112
Adaptation ....... 225
Added value ....... 8
Administrative registers ....... 11
Adult and Lifelong Learning Programme Assessment ....... 210
Advocacy ....... 31, 186
Advocacy coalitions ....... 67
Afghanistan ....... 226
Africa ....... 109, 197, 225, 230
Agencification ....... 24
Agency ....... 18
Agenda setting ....... 38
Agricultural Extension ....... 116
Agriculture ....... 197
Agriculture sector ....... 198
Aid ....... 34
Aid effectiveness ....... 53, 177, 226
Aid policy ....... 181
Analysis ....... 51
Anglophone Caribbean ....... 229
Angola ....... 196
Applied Ethics ....... 213
Appropriation ....... 79
Assessing quality of life ....... 45
Assessment of impacts ....... 142
Assistance for the Modernisation of Farms ....... 6
Audit committee ....... 187

B
Benchmarking ....... 210
Best practice ....... 99
Better Regulation ....... 29
Bibliometric indicators ....... 37
Bio-psychosocial ....... 73
Brain Korea 21 ....... 163
Brazil ....... 91, 92
Broadened assignment ....... 54
Building evaluation capacity ....... 96

C
Cameroon ....... 193
Campaign ....... 186
Capability approach ....... 193
Capacity building ....... 6, 32, 123, 146, 157
Capacity development ....... 17, 71
Capacity development in evaluation ....... 31
Capturing impacts ....... 221
Case studies ....... 35, 110, 189
Case Study ....... 217
Cause-and-effect analysis ....... 204
Cervical cancer control ....... 49
Change ....... 78
Changing evaluation tasks ....... 87
Child and elderly care services ....... 105
Child and family policies ....... 35

Children ....... 80
Children and youth ....... 188
Church ....... 90
Citizen approach ....... 220
Citizen participation ....... 219
Citizens’ rights ....... 131
Citizenship Rights ....... 39
Civic evaluation ....... 141
Civil society ....... 71, 91
Civil Society Organisations ....... 84
Climate Change ....... 186, 225
Cluster analysis ....... 105
CMO logic ....... 104
Co-production ....... 115, 122
Cohesion ....... 137
Cohesion Policy ....... 10, 115, 143, 148
Collaboration ....... 97
Collaborative cultures ....... 96
Collaborative evaluation ....... 49, 109
Commissioner of evaluations ....... 14
Common difficulties ....... 143
Communication ....... 61
Communities of practice ....... 175
Community education ....... 164
Community engagement ....... 164
Community of inquiry ....... 165
Community of practice ....... 146
Community partnerships ....... 123
Community-based Projects ....... 108
Comparative analysis ....... 176
Comparing different delivery systems ....... 155
Comparison ....... 128
Competences ....... 175
Competencies ....... 75, 200
Competition for revenue base ....... 192
Competitiveness ....... 76
Complex evaluation ....... 35, 178
Complex policy ....... 89
Complex policy environment ....... 8
Complex program evaluation ....... 220
Complex systems ....... 13, 202
Complexity ....... 69, 110, 126, 127, 135
Compliance ....... 174
Complicated and complex interventions ....... 118
Comprehensive M&E System ....... 176
Compulsory school ....... 27
Computable General Equilibrium Model ....... 77
Concentration indices ....... 193
Concept mapping ....... 14
Conceptual framework ....... 52
Conference exhibit booth evaluation ....... 222
Conflict ....... 69
Conflict Prevention ....... 134, 178
Conflicting objectives ....... 27
Congo and Sudan ....... 134
Consumer protection ....... 131
Context ....... 173
Context of evaluation ....... 162
Contextual conditions ....... 50
Contextual change ....... 107
Continuous learning ....... 202
Contribution ....... 89
Contribution analysis ....... 12, 13, 204
Control ....... 206
Cooperation ....... 169
Coordination mechanisms ....... 229
Evaluation methodologies . . . . . . 76
Evaluation methodology . . . . . . 126
Evaluation Methods . . . . . . 16, 55, 80, 212
Evaluation models . . . . . . 202
Evaluation of development . . . . . 171
Evaluation of services . . . . . . 188
Evaluation of work services . . . . . 124
Evaluation partnerships . . . . . . 208
Evaluation planning process . . . . . 221
Evaluation plans . . . . . . 101
Evaluation policy change . . . . . . 78
Evaluation practice . . . . . . 222
Evaluation practice in Estonia . . . . . 161
Evaluation process . . . . . . 14
Evaluation public services . . . . . 85
Evaluation purposes . . . . . . 103
Evaluation quality . . . . . . 6
Evaluation questions . . . . . . 113, 126
Evaluation research . . . . . . 35
Evaluation results . . . . . . 125
Evaluation society . . . . . . 143
Evaluation stakeholders . . . . . 208
Evaluation standards . . . . . . 22
Evaluation system . . . . . . 16, 143, 160
Evaluation systems . . . . . . 74
Evaluation theory . . . . . . 33, 117
Evaluation tool . . . . . . 20
Evaluation use . . . . . . 34, 67, 103, 149, 166, 175, 176
Evaluation utilization . . . . . . 14, 67, 78
Evaluation value . . . . . . 47
Evalutional inculturation . . . . . . 22
Evaluative index . . . . . . 211
Evaluative indicators . . . . . . 211
Evaluative thinking . . . . . . 30
Evaluative-thinking . . . . . . 114
Evaluators' Roles . . . . . . 114
Evaluators . . . . . . 101
Evaluators' competencies . . . . . 61
evaluaton . . . . . . 163
Evidence-based management . . . . . . 160
Ex ante impact assessment . . . . . . 198
Ex-ante evaluation . . . . . . 86, 145
Ex-ante evaluations . . . . . . 106
Ex-post evaluation . . . . . . 148, 216
Exit examinations . . . . . . 229
Exploitation . . . . . . 60
External and internal analysis . . . . . . 221
External evaluation . . . . . . 61, 208
External review . . . . . . 30
F
Feasibility . . . . . . 85
Female Genital Mutilation/Cutting . . . . . . 116
Feminist evaluation . . . . . . 173
Fight against discrimination . . . . . . 166
Finnish planning system . . . . . . 86
Fisheries management . . . . . . 199
Fitness check . . . . . . 29
Flemish . . . . . . 44
Focus group . . . . . . 35
Följetforskning . . . . . . 102
Follow-up monitoring and evaluation . . . . . . 229
Formalistic evaluation . . . . . . 187
Formative evaluation . . . . . . 149
Foundations . . . . . . 121
Fourth generation evaluation . . . . . . 190
Frame Analysis . . . . . . 132
Frame analysis approach . . . . . . 220
Framework Programme . . . . . . 125, 144, 216
Funded environmental consulting . . . . . . 199
G
Gender . . . . . . 51, 68, 94, 158, 159, 173, 201, 216, 222
Gender and evaluation . . . . . . 201
Gender equality . . . . . . 158, 219
Gender health policy . . . . . . 93
Gender mainstreaming . . . . . . 201
Gender mainstreaming evaluation . . . . . . 94
Gender policies & programmes . . . . . . 201
Gender policies evaluation . . . . . . 94
Gender-sensitive evaluation . . . . . . 93
Generate evidence on the achievement of results and institutional performance . . . . . . 71
Generating an evaluation culture . . . . . . 93
Ghana . . . . . . 84
Gipuzkoa . . . . . . 223
Global . . . . . . 83, 179
Global health . . . . . . 179
Global health research . . . . . . 179
Good governance . . . . . . 174, 205
Goverenance . . . . . . 24, 144, 151, 156, 221, 225, 230
Government Accountability . . . . . . 91
Government performance . . . . . . 205
Government program evaluation . . . . . . 166
Governments and policy makers . . . . . . 79
Grants for SMEs . . . . . . 10
Grounded theory . . . . . . 168
Group based interventions . . . . . . 191
Growth . . . . . . 10
H
Health . . . . . . 36, 98, 179
Health evaluation . . . . . . 216
Health inequalities . . . . . . 118
Health interventions . . . . . . 153, 154
Health policies . . . . . . 193
Health policy . . . . . . 114, 117
Health program evaluation . . . . . . 29
Health quality . . . . . . 21
Health Sector . . . . . . 82, 99
Health systems . . . . . . 146
Health Technology Assessment . . . . . . 117
Healthcare . . . . . . 141, 200
Heuristics and Knowledge Production . . . . . . 42
High education evaluation . . . . . . 49
High-Tech Strategy . . . . . . 8
HIV/AIDS . . . . . . 92
HIV/AIDS sensitization . . . . . . 65
Home visits . . . . . . 117
Horizontal coordination . . . . . . 115
Housing . . . . . . 118, 227
Human and political factors . . . . . . 95
Human capital . . . . . . 80
Human rights . . . . . . 92, 201, 208
Humanities and social sciences . . . . . . 37
Hungary . . . . . . 143
Hybrid evaluation . . . . . . 114
ICT FP7 . . . . . . 125
IDPs . . . . . . 224
IFI . . . . . . 134
Impact . . . . . . 14, 32, 65, 109, 160, 197, 204, 219, 223, 227
Impact analysis . . . . . . 80
Impact assessment . . . . . . 87
Impact assessment . . . . . . 76, 86, 107, 137, 198
Impact evaluation . . . . . . 12, 35, 39, 77, 116, 118, 138, 150, 153, 154, 198, 212
Impact of evaluation . . . . . . 162
Impacts . . . . . . 18
Implementation . . . . . . 57, 78
Implementation analysis . . . . . . 29
Implementation procedures . . . . . . 119
Implications . . . . . . 219
Incentive structures . . . . . . 135
Independence . . . . . . 47, 89