Perspectives on Impact Evaluation

Approaches to assessing development effectiveness

Conference Abstracts

Cairo, March 29-31
Valuing evaluation beyond programme boundaries: Communicating evaluations to enhance development effectiveness globally.

**Presenter: Johanna Lindstrom**

While much is known about increasing the learning from, and use of, evaluation by a programme’s stakeholders there is often less focus on how evaluations are communicated externally to influence those working in other contexts. Many evaluations never go beyond the programme’s boundaries, or are not easily accessible. However most have knowledge within them which is of relevance globally to a far wider range of development actors: both in terms of the substance and the process of the evaluation.

However, there are a number of challenges and issues specific to evaluations which don’t incentivise or make this an easy task. The session will take as its starting point four issues relating to communicating evaluations:

Communication strategies: How can evaluations be communicated, synthesised, stored and later accessed by those for whom they might be useful? Is publishing in a journal or putting it on a website enough? There are a variety of communication methods and dissemination channels which could be valuable to evaluators such as information and knowledge intermediaries, the broadcast media and international networks. What strategies/channels/methods are effective in communicating evaluations beyond the specific programme context? What kinds of networks and communities could both benefit from and add insight to the final conclusions of an evaluation itself?

Supporting use of evaluations in policy and practice: Whilst it is important to ensure that evaluation results are available, for example through institutional websites, we cannot take it for granted that someone for whom the results of the evaluation are relevant will actually be looking for this type of information, will be able to access it online or will know how to search for it. How can ‘decision-makers’ be encouraged and supported to use evaluations from other contexts/programmes for evaluation informed decision-making?

Context and politics: Knowledge is not a neutral commodity, which can just be picked up and applied elsewhere. The complexity of development processes means that results from evaluations are often context specific and rarely clear-cut. Evaluations are often politically sensitive, and the most useful learning may be watered down into bland lessons learnt which are of little use. How can we ensure context specific and potentially sensitive evaluations are shared, adapted and applied beyond the programme context?

Understanding how evaluation influences: An important step in looking at how to increase uptake of evaluation is to understand how research more generally, and evaluations in particular, influence policy processes and practice. How do we assess the influence of evaluations on policy/practice?
Knowledge of and attitudes towards impact evaluations amongst senior managers in South Africa’s Department of Social Development

Presenters: Tsholofelo Adelekan, Dibolelo Ababio, Dugan Fraser

The high costs of fieldwork and other variables make quantitative impact evaluations very expensive in South and Southern Africa. It is therefore important they are carefully designed and structured so that they meet the needs of their users.

South Africa’s national-level Department of Social Development is primarily a policy making, monitoring and review institution, with implementation and delivery of social development benefits and services in most instances taking place provincially and locally. Monitoring and evaluation is a key part of the national Department’s function and is usually undertaken in partnership with service providers and internal partners, usually including the policy function concerned.

During 2008, working closely with UNICEF and the South African Social Security Agency, which is responsible for the payment of social security grants, the Department has commissioned a large scale, quantitative impact evaluation of the nation’s Child Support Grant, a five-phase study to be undertaken over four years.

During our interactions with fellow staff members in the policy functions we support in the M&E unit, both as we prepared the terms of reference and made other arrangements for the evaluation, we have become increasingly aware of the extent to which the term “impact evaluation” is used casually and interchangeably with other terms such as review, study, assessment and appraisal and, furthermore, of the extent to which expectations of what an impact evaluation can achieve are both unrealistically optimistic while also being pessimistically minimal.

We have resolved that it is important that we get a better understanding of what our colleagues mean when they say “impact evaluation”, so that we can (i) try to ensure that what we deliver at the end of an arduous process is useful and meets their expectations; and (ii) simultaneously manage expectations and make them more realistic and better matched to what will be produced. In Patton’s “Utilisation Focused Evaluation” terms, we seek to understand both the users of our evaluation better and the uses they seek to make of it.

In order to understand our colleagues knowledge and understanding of impact evaluations we will be making use of a semi-structured self completed questionnaire distributed to around 40 senior managers and we will follow these up with in-depth interviews with around 15 respondents. Both the interviews and the questionnaires will explore their knowledge and expectations of impact evaluations. Preliminary results will be shared with conference participants.
Cultivating Demand within USAID for Impact Evaluations of Democracy and Governance Programs

**Presenters: Margaret Sarle, Devra Moehler, Mark Billera**

There are countless testimonies by those who credit USAID-funded programs with improving conditions in their lives and in their countries. Gathering quantifiable evidence as proof of the effectiveness of USAID assistance has been more elusive. This can be especially challenging when it comes to measuring the positive impact of USAID assistance addressing certain aspects of democracy, such as civil society, good governance, and respect for human rights.

Recently, USAID’s Office of Democracy and Governance (DG) turned for guidance to the nonprofit National Academy of Sciences and its affiliate, the National Research Council (NRC). These were formed by charter by the US Congress to provide independent policy recommendations to the government. The NRC convened a group of leading experts in the fields of democracy promotion and evaluation to examine current USAID practices and to provide recommendations for improving how USAID evaluates the effectiveness of its democracy assistance programs.

The NRC report entitled, “Improving Democracy Assistance: Building Knowledge Through Evaluations and Research”, acknowledges the extensive efforts already taken to monitor and evaluate the impact of democracy promotion programs, but includes a set of specific recommendations for improving those efforts. The central recommendation is to undertake a pilot program of impact evaluations designed to demonstrate whether such evaluations can help USAID determine the effects of its DG projects on targeted policy-relevant outcomes. A portion of these impact evaluations should use randomized designs since, where applicable and feasible, they are the designs most likely to lead to reliable and valid results in determining project effects and because their use in DG projects has been limited.

Two main challenges now confront USAID’s DG Office. First, we must figure out the best way to undertake impact evaluations of our DG programs, both technically and bureaucratically. Second, we must convince our USAID and implementing partner colleagues around the world that such evaluations will be worth the time, effort, and possible additional costs. The presenter(s) will discuss the steps that USAID is taking to meet these challenges and how good answers to the first challenge are crucial to overcoming the second one. Examples from Africa will be included.

A Syndrome of Irregular Enthusiasm (SIE): Increasing the utilisation of evaluation findings in the UPHOLD Project

**Presenter: Apollo Nkwake. M**

The paper explores the reason why after undertaking costly impact evaluation activities, organizations are not enthusiastic enough to implement findings of and recommendations from such studies. The paper draws on experiences of the Uganda Program for Human and Holistic Development (UPHOLD) in attempting to suggest an explanation and panacea to this phenomenon to be rooted in the development of an evaluation culture, and a model for evaluation proceeds usability is suggested.
Evaluation participative des indicateurs de conservation et de bien être dans le Paysage forestier du Bassin du Congo

Presenters: Dominique Endamana, Louis Defo, Antoine Eyebe, Jeffrey Sayer, Manuel R. Perez, Intu A. Boedihhartono, Cléto Ndikumagenge, Kamiss Ami

Le paysage forestier tropical du Bassin du Congo est le deuxième massif forestier du monde après celui de l’Amazonie. Il couvre une superficie de plus de 228 millions d’hectares et est riche en ressources faunique, floristique, halieutique et hydrique. Depuis 2004, les organisations internationales de conservation et de recherche ont porté le processus de suivi – évaluation des changements survenus dans le paysage de ce complexe forestiers à travers 34 indicateurs sélectionnés de manière participative. Ces indicateurs, regroupés en 5 groupes d’acquis capitaux: humain, social, naturel global, naturel local et physique, ont fait l’objet d’une cotation de 1 (situation pire) à 5 (situation souhaité) appelé échelle de Likert. Cet article montre la tendance évolutive des indicateurs de conservation et du développement dans le Tri National de la Sangha qui comprend le Cameroun, le Congo et la République Centrafricaine. Les résultats du suivi entre 2006 et 2008 montrent deux principales tendances: interchangeabilité entre les acquis de conservation et ceux du développement; conflit entre les acquis de naturels globaux et locaux sur le développement socioéconomique. L’étude recommande la poursuite du suivi des indicateurs ce qui permettra à moyen terme de tirer des conclusions précises, identifier de meilleures opportunité, assurer de bon compromis et influencer les politiques pour une meilleure conciliation de deux paradigmes conservation - du développement dans les projets/programmes de conservation.

Evaluation Results: Engaging various audiences for effective learning and influence

Presenter: Jerushah Rangasami

This paper presents two important challenges faced by evaluators with regards to achieving influence through evaluation and presents possible solutions to these challenges using examples from recent evaluation practice.

The first challenge concerns the difficulties of effectively reaching a wide audience with evaluation reports. Often, as producers of evaluation reports that are sometimes long and fairly technical, we hold unrealistic expectations of our intended readership, especially when this readership comprises of multiple and diverse stakeholders. While it is not feasible to dilute the integrity of an evaluation report for the sake of readability or to enable dissemination, it is possible to produce a separate summary report that is easily accessible to stakeholders who may otherwise not engage with the material. “Accessibility” can take the form of clear and professional design, the use of diagrams and pictures and the style of writing. Our experience has shown that these accessible summary reports extend the reach of evaluation findings far beyond a standard evaluation report, towards influencing policy at multiple levels. In addition, if carefully written with the relevant audiences in mind, summary reports can be disseminated during workshops (preferably with various groups of stakeholders present) at which findings are presented, copies are distributed and discussion is facilitated.
Summary reports can be planned after the full report is produced, but another approach to help ensure effective reach is to consider the report audience(s) in the evaluation planning stage. Evaluators can do this in a participatory way by including a session on reporting or ‘audience’ in an evaluation planning workshop with key stakeholders. In this session, participants identify key audiences for the report and develop a strategy with the evaluators to ensure that an accessible report is developed for these audiences, usually in addition to the full report. Another consideration is which groups of stakeholders should get which report – a full report that is comprehensive but is not likely to be read, or a summary report that is not comprehensive (but references the complete report) but is more likely to be read. The second challenge presented here is the lack of meaningful engagement with evaluation results by the commissioning or implementing organisations. Evaluation findings can end up merely as reports on desks, or as a once-off, one-way presentation by the evaluator. Without a forum in which to discuss and debate them, evaluations can become a perfunctory fulfilment of a contractual duty for stakeholders, rather than an opportunity to engage in constructive programme development and to inform future strategy. Once again, experience has revealed that a highly participative evaluation findings workshop with a range of key programme stakeholders results in active engagement with evaluation findings. A forum such as this provides the impetus for civil society engagement and programme development towards sustainability. As mentioned above, summary reports make findings more accessible and stakeholders are much more likely to participate in such a forum if the findings are initially accessible.

Using Knowledge Management functions to achieve influence

**Presenter: Amuzweni Lerato Ngoma**

The South African government created youth development machinery that aims to ensure that the socio-economic and political needs are taken into account across all sectors of society, part of which comprises the Umsobomvu Youth Fund (UYF). The UYF is mandated with the implementation of programmes that develop the skills of young people and facilitate their employability and employment. The UYF has been implementing programmes for 7 years, wherein a lot of experience has been generated in the pursuit the youth development mandate. How has this knowledge on the impact of the institutions been created? This paper focuses on how this knowledge has been area by the Umsobomvu Youth Fund.

Within the UYF, various institutional processes, systems and tools have been designed to guide the commissioning, doing and management of impact evaluations. However, the Monitoring and Evaluation function is separate from the core-businesses of the organisation, in which it is a support function. This has resulted in the M&E function and its outcomes being seen as an outside perspective and requirement. The ultimate result of this situation is that the outcomes of impact evaluation, for example, in terms of the information and knowledge generated there-from become compromised in terms of how the outcomes are used.

In this situation, impact evaluators, which are typically outsourced to consultants in order to maintain objectivity, have to enter the organisation, and inform the M&E practitioners, together with the representatives of whatever intervention. The impact evaluators have the duty, in this context, of almost educating these stakeholders on what the experience of implementation was, what was achieved and what stands to be done with the intervention going forward.

It is in this context that knowledge management can be a key tool in the institutionalisation of the impact evaluation outcomes and knowledge.
This paper details a good practice in communicating impact evaluation findings. It argues that the best way is to return to the owners and actors of development programmes and communicating with them the key findings. This communication, which was used in a programme close-out process, is best conducted by coordinating people with the same interests and hosting a learning forum or workshop. The key findings from the impact assessment are been documented and synthesised before the workshop, as well as communicated same-interest groups before the learning forum. At the learning forum, new knowledge around the findings was sought-out. Participants spoke about how they would use the findings to improve their practice. This new knowledge and information should be geared towards informing how best the findings can be integrated for the improvement of programme practice. The entire process was thus communicated with the UYF's executive committee.

**Evaluating Advocacy and Influence Impacts to Increase Development Effectiveness**

**Presenter:** Sanjeev Khagram

**Co-authors:** Julia Coffman, Subarna Mathes, Ricardo Wilson-Grau

The complex and dynamic nature of advocacy programs makes impact evaluation of such efforts difficult and yet all the more necessary when addressing development effectiveness. It compels practitioners to think about existing evaluation methods in different ways. Due to this demand, innovative approaches are being developed and next generation good practices are being identified to examine, measure, assess and learn from the governance, institutional and policy impacts of advocacy efforts.

Despite its relative youth, the field of advocacy impact evaluation is extremely dynamic and quickly gaining momentum through the combined efforts of evaluators, academics, funders, and advocates. Alongside the development of new and adapted tools and methods is also the emergence of common advocacy evaluation norms and principles. While there are no hard and fast rules for evaluation practice in this arena, these developments are beginning to identify evaluation approaches that make the most sense in the advocacy context. This session will provide an introduction to innovative approaches and next generation good practices to examine, measure, assess and learn from the impacts of advocacy efforts, especially when addressing development effectiveness. The work is based on a Gates Foundation funded project that examined innovative methods and cases of advocacy impact, learning and evaluation.

**The application of the emerging Government-wide Monitoring and Evaluation System for policy impact evaluation in South Africa**

**Presenter:** Fanie Cloete

The evidence-based policy analysis approach is an increasingly popular paradigm for public policy assessments. It is still in an emerging state in South Africa. Until recently, no coherent system of systematic policy monitoring and evaluation existed in the South African Public Service. This state of affairs, however, changed in 2006. The Presidency is currently coordinating the roll-out of a massive implementation programme known as the Government-
Wide Monitoring and Evaluation System (GWM&ES). This programme is intended to establish a uniform system of monitoring and evaluation across all spheres of government and also including the business and voluntary sectors of South African society. If the implementation of this programme is successful it will have huge potential for the improvement of public policy results in this country. Successful policy implementation, however, depends on doing the right things right. The system is still emerging and struggling to consolidate its new role across the governmental sector. The SA government has further acknowledged the absence of a clearly formulated and generally accepted national vision that seriously complicates the need to move beyond its current focus on policy output evaluation to medium and long term policy outcome and impact evaluation in order to create a better life for all. The question is what contribution the emerging GWM&ES will be able to make to improve this situation.

This paper will firstly summarise the main elements and general requirements of the evidence-based approach to policy analysis and assessment. It will then summarise the emerging properties of the GWM&ES as a complex adaptive system (CAS) and critically assess the appropriateness of the emerging GWM&ES to achieve the SA government’s policy goals. Positive and negative trends in the development of the GWM&ES will be identified and assessed. The paper concludes with a summary of recommended strategies for optimal implementation of the GWM&ES in order to achieve maximum efficiency and effectiveness for purposes of evidence-based policy impact assessment. Strategies to overcome existing obstacles in the successful roll-out of the system for this purpose will also be summarised.

### Theme 4: Achieving influence: reporting, outreach, learning and policy impact

<table>
<thead>
<tr>
<th>Session number: [1.4]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Session stream:</strong> Achieving influence</td>
</tr>
<tr>
<td><strong>Session name:</strong> Achieving influence</td>
</tr>
</tbody>
</table>

**Impact Evaluation in Practice: A Review of Impact Evaluations conducted in South Africa over the past 5 Years.**

**Presenter: Benita van Wyk**

This paper summarizes & reviews a convenience sample of so called “Impact Evaluations” commissioned by selected government departments / agencies in South Africa over the past 5 years. The purpose of this research is to explore the understanding of the concept “Impact Evaluation” as it is applied in the South African context. This practical understanding of Impact evaluation as it is implemented on the ground is contrasted with various theoretical understandings of impact evaluation.

The review focuses on:

- The kinds of questions asked (We use the typology suggested by Chelimsky (2007). Descriptive, Normative, Attributive, Analytic-Interpretive),
- The unit of analysis in impact evaluations (Policies, Programmes, Projects, partnerships, institutions, networks, etc.),
- The timing of impact evaluations in the implementation / organisational cycle,
- The research designs & methods employed (Experimental, Quasi-Experimental, Mixed-Methods, Qualitative Methods, etc.),
- The kinds of learning supported by the conclusions and recommendations from impact evaluations (process learning, organisational learning, impact learning, knowledge development and policy learning), and
- The intended use supported by impact evaluation - We refer to use as discussed by Marra (2000), Patton (1997), Sandison (2006) and Weiss (1999).
It identifies the most common type of questions answered in impact evaluations and reflects on the limited number of impact evaluations that manage to answer attributive questions. It identifies the different units of analysis when impact evaluations are conducted and reflects on the timing of impact evaluations in relation to the project, programme, institution or project timeline. It identifies a variety of methods and evaluation designs employed in the South African context and comments about the appropriateness of the methods / designs to the impact evaluation questions. Finally it identifies context specific uses of the term “impact evaluation” and points out some possible gaps between the theoretical concept of impact evaluation and the practical application thereof.

<table>
<thead>
<tr>
<th>Theme 4: Achieving influence: reporting, outreach, learning and policy impact</th>
<th>Session number: [1.4 ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session stream: Achieving influence</td>
<td>Session name: Achieving influence</td>
</tr>
</tbody>
</table>

**Program evaluation experience in Uruguay (2005 - 2008): a mixed-method approach**

**Presenter: Ignacio Pardo**

The impact evaluation of a social plan (involving several programs) in Uruguay (from 2005 to 2007/2008) was designed combining both quantitative and qualitative techniques and therefore innovating through a mixed-method approach. Better and more comprehensive impact evaluation designs can arise in the near future from the lessons learned with this experience.

From 2005 to 2007/2008, the Ministry of Social Development (Ministerio de Desarrollo Social) tried to address the social emergency triggered by the economic crisis of 2002 in Uruguay with a plan called PANES (Spanish acronym for Plan Nacional de Atención a la Emergencia Social), aimed at the poorest households. This plan involved several programs, the most important of which include: an income transfer program ("Ingreso Ciudadano"); a community work program administrated by NGOs, to reintegrate participants into the labor market ("Trabajo por Uruguay"); and "Rutas de Salida", which was the PANES component that tried to provide exit strategies for the participants.

The evaluation of the impact of the program was based on a mixed-method approach, combining both quantitative and qualitative data from techniques such as surveys, focal groups and in-depth interviews. The design of the study was flexible and analysis was focused on the interrelation between the two types of data. Through the study, the most relevant dimension taken into account was the voice of the participants.

The main lessons learned from this study include the correct use of focal groups and in-depth interviews when working with marginalized communities; the different ways of measuring empowerment strategies; the reflection on the consistency of interpretations across different actors (NGOs, participants, political authorities, evaluation experts); and the importance of being aware of design limitations in order to improve the quality of evaluation research.
Proving to policy-makers that we have an impact on reducing under-nutrition in Ethiopia: Save the Children UK’s Applied Learning Information System

Presenter: Mathew Hobson

To address malnutrition in Ethiopia, a stronger evidence base is needed around a range of questions such as “what makes the difference to children’s nutrition status – health, wealth, knowledge, or a combination of the three?”. Building this evidence base is essential for supporting advocacy and for improving the effectiveness of programming, in Ethiopia and beyond.

Save the Children UK’s Hunger Reduction Thematic Programme Plan (HRTPP) consolidates livelihoods, food security and nutrition activities under one umbrella framework. The HRTPP is a strategic and pro-active approach to impacting hunger in Ethiopia. The pillars of the HRTPP are organized around the UNICEF conceptual model of the determinants of malnutrition. The model tackles the intermediate determinants of child malnutrition by directly improving “wealth” (livelihood opportunities, income, and assets), and “knowledge” (caring practices) and by indirectly affecting “health” (e.g. through improving financial access to health care). Livelihoods and Nutrition Information Systems will support development programming decisions and contribute evidence for advocacy purposes.

Why and how is the Applied Learning Information System innovative? The conventional approach to project M&E and impact considers (mainly) output indicators and assesses the impact of each project on its own merits. Conventional M&E of projects does not consider impact on organizational strategic goals and objectives (only project goals/objectives); comparative effectiveness/efficiency between health, wealth and knowledge interventions; comparative cost-effectiveness of health, wealth & knowledge interventions; impact of any capacity building work or impact of any advocacy work. The HRTPP will add value over individual project M&E approaches, by:

- organizing activities around a unified conceptual model of malnutrition reduction,
- incorporating capacity-building, partnership, and advocacy objectives that hope to impact malnutrition through non-project channels,
- enabling a more cohesive and strategic approach to leveraging resources for these activities.

With a strategic approach to addressing hunger that goes beyond the sum of individual projects, the HRTPP is also uniquely positioned to plan and carry out an applied research agenda to ask, and answer, questions that cut across projects and go beyond ‘typical’ project M&E.

Because the HRTPP is an umbrella for many independently operating projects (there are 17 Save the Children projects currently operating to reduce malnutrition in Ethiopia), there are several study design challenges in this cross-project learning model that do not typically apply to single project evaluations. Impact Monitoring is the core research component of the ALIS, and will require the most time and financial resources. In order to produce plausible results, SCUK will employ the most widely accepted quasi-experimental study design used in population-based research (Habicht 1999). Participatory Rural Approaches and Tools will be used to unpick the causes of changes at household level, as well as identify the type of participation of communities according to Types of Participation (Pretty and Cornwall 1996).
Use of impact evaluation for organizational learning and policy influence: The case of international agricultural research

Presenters: Jamie Watts, Douglas Horton, Ronald Mackay, Patricia Rogers, Sheelagh O’Reilly

In the context of international agricultural research, most impact evaluation work has focused on assessing the adoption of new varieties, increases in yields, and rates of return to research investments, where the establishment of a counterfactual is relatively straightforward. However, a large proportion -- perhaps 75% -- of the work now being done under the banner of agricultural research is in the areas of natural resource management, policy and institutional research, germplasm conservation, and capacity development – areas that present challenges for traditional impact assessment. Work in these areas is characterized by collaborative knowledge generation, complex innovation processes, and adaptive management. The move towards more collaborative agricultural research for development is similar to the development of "translational research" in the health field, where it is recognized that uptake and use of evidence from research needs active and informed translation to specific practice contexts, and that evidence from practice is also needed to inform research.

Traditional impact evaluations in the field of international agricultural research have had limited impacts on policies and on organizational learning, and there is increasing pressure to improve the quality and use of evaluations. The presentations in this session will outline the challenges to use of impact evaluations in historical perspective, and will describe some avenues currently being pursued to broaden the range of evaluation methodologies and approaches to address these challenges, with the ultimate goal of improving the utilization of impact evaluations for learning and policy influence. The presentations will be the following:

- Presentation 1: Use and non-use of impact evaluation: the CGIAR case
- Presentation 2: Towards a broader range of impact evaluation methods (BRIEM) for collaborative research: report on a work in progress
- Presentation 3: Role of Impact Evaluation in Moving from Research into Use
in processing better-quality food, delivering it to consumers, and ensuring that people can get access to the food they need to live healthy and productive lives.

But with the recent food and financial crises still weighing heavily on the international development agenda, there is an urgent need to learn more about "what works" in agriculture—what sorts of policies, programs, and investments in agricultural development have actually improved food security and reduced hunger. This paper sets forth a framework for identifying, documenting, and analyzing evidence on what works in agriculture so that cross-cutting lessons from past successes can be learned, shared, and applied to future investments in agricultural development. This framework relies significantly on data, information, and analyses provided by rigorous impact assessments of interventions in agricultural development.

Note that this paper is part of an 18-month initiative designed to identify, document, analyze, and communicate lessons learned from proven successes in developing-country agriculture. Although this initiative is still in its initial stages, a complete set of findings will be produced by June 2009. Thus, this paper, which aims to provide the conceptual and methodological groundwork for the overall initiative, is a work in progress.

This paper addresses several issues that are specific to proving success, i.e., to demonstrating real impact with rigorous tools and methods of analysis. First, given the complex nature of food security, this paper aims to develop a clear understanding of the terms used to describe its causes, possible remedies, and pathways through which these remedies operate. Second, given that thousands of interventions may claim to have succeeded in addressing food security, this paper develops clear selection criteria to identify and select “proven successes” in a rigorous and transparent manner. Finally, this paper develops a set of definitive indicators to demonstrate how a given intervention has had a demonstrable impact on food security, and how comparisons can be made across interventions.

Ultimately, this paper provides a set of solutions that will help to identify, document and analyze lessons learned from proven successes in developing-country agriculture. These solutions all revolve around data and analysis provided by rigorous impact evaluations, thus demonstrating the importance of institutionalizing impact evaluation for agricultural development, and communicating the findings of impact evaluations, to ensure the efficient allocation of scarce resources for development.

### Theme 2: Impact evaluation: Approaches and methods

### Theme 6: Impact evaluations of policies, programs, projects, partnerships and networks

**Session number: [2.1 ]**

**Session stream: Agriculture**

**Session name: Agriculture**

---

**Initiation of innovation platforms in small scale agriculture in Lake Kivu: Issues, challenges and lessons learnt for impact evaluation**

*Presenter: Pamela Pali*


Innovation platforms present a unique model in reversing the declining African agriculture. This builds on a premise that through research, agriculture specifically in Sub Saharan Africa has registered some successes. However, these successes have not had wide scale impact due to (i) Poor linkages between production, processing, trade/marketing and consumption (ii) Inadequate human & financial resources and (iii) Weak institutional frameworks including partnerships for
addressing these issues. Success stories have demonstrated the benefits of linking production to markets and the role of policies and institutions in assuring success. This underscores the benefits of innovation (social process) that catalyze the generation and implementation of combinations of knowledge from multiple sources to create development impact. Within this context, the building blocks essential for impact evaluation are in the process of being developed. Monitoring and evaluation of innovation platforms shall play a critical role in the final assessment of impact evaluation. The paper assesses the dynamics involved in establishment of IPs under a range of different contexts including locations, policy regimes, farmer organization and competences of stakeholders. It outlines the process IP establishment in Rwanda, Uganda and the democratic republic of Congo, teasing out the key issues, challenges and lessons learnt. This paper does not challenge the importance of investing in science and technology capacity, which is well recognized in innovation systems theory. Rather it focuses on the additional insights of harnessing mutual partnerships among stakeholders and types of interventions that can be derived from an innovation systems perspective and that can influence the generation and use of science and technology for wider impact.

**Theme 2: Impact evaluation: Approaches and methods**

**Theme 6: Impact evaluations of policies, programs, projects, partnerships and networks**

**Session number: [2.2]**

**Session stream: Agriculture**

**Session name: Agriculture in Africa 1**

**NONIE Hands on Clinic: impact evaluation options discussed around a real planned IE in agriculture**

**Presenter: Paul Gertler**

NONIE, though its mission – doing more and better impact evaluations – strives to fill existing knowledge gaps of what works in development by focusing its work on a particular set of themes. Capitalizing on the successes in its October 2008 meeting in Lisbon, NONIE continues its learning-by-doing agenda by organizing impact evaluation design clinics. The two sessions build on the first Conference session proposed by NONIE (More and better impact evaluation for learning “What Works in X” – Problem solving session to address knowledge gaps) with two exclusively interactive clinics in the same thematic focus area. These two clinics invite specialists to work together with participants to provide expert comments and peer review advice to improve the quality of impact evaluations currently in the planning stages.

The clinics start with a presentation of an impact evaluation of an actual project currently in the early planning stages. Led by a moderator, participants work together hands-on with the technical expert and policy reviewer to develop the design and implementation, plan data collection options, and propose analytical approaches for strengthening the impact evaluation. Benefiting from advice and expertise from the technical expert as well as concrete implementation considerations from the policy reviewer the clinic participants will learn options, challenges and trade-offs in planning a real impact evaluation in these highly interactive sessions.
The need to increase productivity enhancing technologies in agriculture through increased investments in agricultural research has grown in importance. Given the need to increase investments in agriculture, it follows that impact assessment studies have become handy as funding agencies (donors, government and private sector) demand better accountability for their funds and empirical justification for further investment.

Experts have also reasoned that increasing resources for agricultural technology is an indispensable but far from sufficient condition for Africa to resolve its low productivity problems. The quantity of resources is as much important as the quality of spending so that resources are used efficiently and more could be achieved with the same level of resources. The lack of an effective strategy and basis on how investments in agriculture could improve productivity in agriculture and thus eradicate hunger is the problem for many African countries.

A study is hence underway which aims to illustrate the use of the Dynamic Research Evaluation for Management (DREAM) model to assess the impact of investing in agriculture technologies and to set priorities for crop based research activities in Zambia. Using secondary data from the Ministry of Agriculture and other sources, the study hypothesises that different priorities for agricultural technology development will result in different patterns of investment; and consequently different growth effects on the agriculture sector and the entire economy. DREAM is a computer program conceptually based on the economic surplus theory and designed for research priority setting and ex-ante evaluations. It computes the net present value of benefits to both producers and consumers as a result of investing in agriculture technology.

Preliminary findings reveal that investing in maize technology (such as improved seed) gives the highest return to smallholder producers, and tobacco for large scale producers. Maize also yielded the highest returns for consumers in Zambia. In order of priority for research investment this entails that maize should be given the highest priority followed by tobacco, cotton, soya bean, groundnuts and millet. The largest binding constraint to agricultural growth in Zambia is the dependence on rain fed production and poor physical and marketing infrastructure. It is essential that as government and other stakeholders focus on raising agriculture productivity through science based technologies, complementary infrastructure for irrigation, and other physical infrastructure as well as supporting institutions be developed in order to achieve the targeted growth in agriculture and the economy as a whole.
The Impact of NERICA rice varieties on Productivity: Evidence from rice farmers in Guinea and Who Benefit More From NERICA Varieties? Gender Impact on Production And Income In Benin

**Presenter: Florent M. Kinkingninhoun-Medagbe**

It is believed that an improvement in crop productivity based on new varieties is a foundation for the potential Green Revolution in developing countries. The Africa Rice Centre developed the New Rice for Africa (NERICA) rice varieties that are said to provide hope for raising the productivity of upland rain-fed rice farming in Africa because of their reported high yield potential and adaptability to the African conditions.

The NERICA varieties were introduced to rice farmers in Guinea through Participatory Varietal Selection (PVS) and on-farm research trials in 1997. The introduction of improved rain fed varieties of rice was initiated by the National Office for Rural Development and Extension (SNPRV), and the non-governmental organization Sasakawa Global 2000, and supported by the local mission of the World Bank and the Special Program for African Agricultural Research (SPAAR). Using data from farm households in Guinea, this study aims at shedding some light on the impact of NERICA adoption on rice yield. We apply the counterfactual outcomes framework to estimate the Average treatment Effect (ATE) of NERICA adoption on yield using semi-parametric and parametric models. The results of the analysis indicate that NERICA adoption significantly increases rice yield by 500 kilograms per hectare suggesting that there is scope for improving rice production in Guinea through a wider adoption of NERICA rice varieties.

This paper examines the gendered impact of NERICA (New Rice for Africa) adoption on yield and income using data from 304 rice farmers randomly selected from Benin. The NERICA® are modern rice varieties produced through WARDA’s interspecific hybridization breeding program, which won its creator Monty Jones the 2004 World Food Prize. They are widely believed to offer hope for Africa’s Green Revolution because of their ability growth under multiple stresses as well as their high response rate to inorganic fertilizers and other inputs. Many studies argued that any change in farm systems affects men and women differently. This is partly due to the differences in perception regarding technology that exist between women and men in farm households. This paper applies the counterfactual outcomes framework of modern evaluation theory using the Local Average Response Function (LARF) to estimate the Local Average Treatment Effect (LATE) of NERICA adoption on farmers’ yield and income. The results show that NERICA adoption has positive and significant impact on farmers’ yield and income. Moreover, the impacts of NERICA adoption are higher among women farmers than among men farmers. The findings suggest the extension of NERICA® varieties in Benin and in other African countries could enhance productivity and welfare. Moreover, targeting women with NERICA can significantly increase rice productivity and consequently total production and income, more than targeting men with NERICA®.
Can Sasakawa for Africa learn and become even more effective, through Impact Evaluation?

Presenter: Roberto La Rovere

Yes it can – and it is doing so. Also through impact evaluation; but it wasn’t easy... It’s out of the question that the Sasakawa Africa Association / Sasakawa Global 2000 (SAA) has been having impacts in Africa over the past 2 decades, and is very often cited by African leaders as a model for promoting agricultural growth in Africa. Yet they are considered by many as an organization that adopts a rather old fashioned, rigid top down model of development that may not always meet recent challenges. Additionally, the organization has been at times defined by knowledgeable observers as being ‘unable to learn’ from its shortcomings as well as successes. SAA also used to lack a systematic way to document and quantify its impact objectively besides anecdotal success stories and few evaluations. As a result, one of SAA main donors, the Nippon Foundation, took a landmark initiative in 2005 to invest in an innovative, extensive, evidence-based external monitoring and assessment system commissioned to CIMMYT for the period 2006-2010, in coincidence with the current phase of SAA investments (see: http://sg2000ia.cimmyt.org – website of the “Knowledge and Monitoring system to Assess Impacts of activities promoted by SAA and partners in Ethiopia and Uganda”: ‘K&MIA’).

We describe the nature of the system implemented, the challenges faced by the external team of scientists from CIMMYT and partners and local institutions in assessing impacts, the solutions found we progressively, the methodological and management challenges we encountered and the ways they were solved to implement such comprehensive K&MIA with limited resources. We distill the lessons we learned in implementing such a mechanism of evidence-based learning and institutionalizing the approach in SAA within the day-to-day challenges typical of Africa. After 3 years of data collection and a two-way feedback between the CIMMYT/K&MIA teams and SAA, at least 2 examples of change catalyzed by the K&MIA have taken place already: 1) SAA is changing its ways and strategies for investing in Ethiopia, to better target those farmers who are actually more in need - and in the more appropriate areas, discontinued activities, and concentrated in fewer hot spots its otherwise thinly spread investments, and 2) the One-Stop Centers (OSCA) approach in Uganda, often mentioned as a success story with not much critical questioning, has been undergoing a deep review of its cost-effectiveness and efficiency in reaching the targets as opposed to alternative providers, e.g. the private sector, and is under review. The paper also discusses the essential communication steps that from now will be used to bring the approach to the attention of investors in Africa, not last the Gates Foundation, among those interested in investing in extension through SAA. The paper makes use of lessons learned and of a synthesis of 3 papers for the 2009 IAAE Conference, that demonstrate the link between more effective extension services and increases in adoption and yields for participants, hence providing insights for the K&MIA and SAA in moving now into policy advocacy.
Systemic Impact Evaluation: a method for evaluating the impact of complex development projects

Presenters: Jean-David Naudet, Jocelyne Delarue, Tanguy Bernard

Commonly used impact evaluation methods rely on quantitative approaches, experimental or quasi-experimental. These methods are particularly efficient to prove and to measure the impact of a simple treatment but they rarely address the needs for a more detailed understanding of why the treatment has worked or not. Applied to international development projects or programs, they cannot always be used. When they are, they often fail to capture the complexity of the intervention and what they prove is not always easy to interpret. Many development practitioners ask for a combination of qualitative and quantitative approaches in impact evaluations, but the academic teams in charge still tend to be too specialised in one method to be able to ensure this combination.

This paper presents step-by-step a systemic methodology using a qualitative approach to assess the impact of complex development projects. It presents its application on an agricultural development project in Guinea, where the outcome of interest is the farmers’ income. The methodology that is presented relies on two well-known methods.

First, it uses the cost-benefit analysis techniques to quantify the differential between the project and its counterfactual. Usually used ex ante, the economic model in this case is built on the basis of data collected on the population concerned and not concerned by the project. The cost-benefit analysis is carried out at the farm level and includes data of the year and recall data. Moreover, it combines an ex post impact assessment and an ex ante model.

Second, the method uses the concepts of farming system analysis to strengthen the internal validity of the ex post impact evaluation. At the regional level, a systemic contextual analysis allows choosing a village which makes a proper counterfactual for the project area: the choice is based on the in-depth understanding of development trends existing in the region prior to the project. At village level, farming system analysis helps to choose in both situations (with and without project) the observation units which were similar before the project and which would have evolved alike in the project’s absence.

Triangulation based on secondary data and direct field observations is very often used in order to corroborate recall data and present interviews and to make the impact evaluation more robust. Applying this method allows identifying the direct and indirect effects of the project, establishing a with and a without project scenario for different types of beneficiaries and quantifying the impact indicator chosen. For each type of farming system, not only does it give a comprehensive model of the adoption of innovations but it also describes the trade-offs the farmers had to accept in order to integrate new crops and why they had to abandon others, emphasizing the fact that benefits were not entirely additional.

This approach moreover proffers a great deal of reliable qualitative information likely to orient the project’s decision-makers and guide the elaboration of new interventions.
Impact assessment by farmers a complementary tool to agronomic / technical and socio-economic evaluation in Tanzania

Presenter: Elizabeth Maeda

The participation of farmers in impact evaluation of technology dissemination has had wide implications on creating a social relationship in which the researchers, extensionists and farmers are all active participants. The active participation of farmers in diagnosis, designing, planning, implementation, dissemination and adoption has been helpful in relations to evaluation leading to an understanding of farmers objectives, concepts, needs, problems and opportunities.

A number of studies on participatory impact assessment have been carried out and research problems have been selected with adequate field validation on their likely productive or economic importance in the past while the target area was selected because of the gap between actual (farmer’s yields) and potential yield appears to be large when compared with the national average. The major factors limiting increased yield along with agro-ecology, are poor agronomical practices due to land degradation, inadequate tillage and seed-bed preparation, inappropriate fertilizer application, rates and types, use of varieties with low yield potential and poor environmental adaptation, and insufficient weed, insect, fungal diseases and socio-economic in nature example land fragmentation.

The study was conducted for five years along the highlands of Northern Tanzania in two villages involving 360 villagers living around the watershed catchments areas. The main objectives of the study was to assess and evaluate several technologies generated on-station or on-farm at different stages of their development, whether during technology adoption on-farm at technology verification which was also done on-farm. A total of 360 randomly selected farmers (both from informal and formal survey) were visited to determine whether or not the population was homogenous. The study area and its population appear to be largely homogenous in terms of climate, soil type, crop pattern, cultivation practices, family composition, institutional support, family size, land tenure, capital assets and existing technology and constraints. The sample size was kept high so that variations between households could be recorded.

The study focused on farm-level data collection assessment and evaluation and analysis of the qualitative data on the socioeconomic structure of households and farms, farm parameters, enterprise patterns, production practices, technical production problems, marketing and consumption patterns. Farmers were involved in the evaluation at three different stages in the screening process that involved evaluation of a large number of alternative prototype solutions to farmer’s problems. These were screened selectively to identify the most promising options from a large number of options and farmers were only exposed to a few well-developed alternatives that appear promising from the researcher’s point of view.

The results obtained from the study provided information on how farmers weighed agronomic/technical, economic, ecological and socio-cultural considerations to arrive to their own conclusions about the usefulness of a new technology in their particular farming circumstances. Increasing the participation of farmers in evaluation of technologies has resulted in high adoption rate of the recommended production practices among farmers in the area that has in turn resulted in increase productivity.
Randomised Controlled Trials in practice: Impact assessment of Helvetas' organic cotton projects in Burkina Faso

Presenter: Peter Schmidt

Helvetas, Swiss Association for International Cooperation, is a development NGO working in 15 countries of the South and East with an annual budget of around 50 Million USD. Helvetas’ working areas include "Rural Infrastructure", "Natural Resources Management", "Education and Culture" as well as "Civil Society and State". Helvetas implements around 170 grass-root oriented projects. Representative for its thematic and geographic portfolio Helvetas has selected ten "flagship projects" for close impact monitoring and learning, among them a project to promote the introduction of organic cotton production and marketing in Burkina Faso as an example for the working area "Natural Resources Management". Using randomised controlled trials the impact study looks into socio-economic and ecological aspects in purposively selected villages. It compares the socio-economic situation of randomly selected farm households participating in the programme – these are organically producing farms – with comparable farms, which still produce conventionally. This quantitative part of the impact study is complemented with focus group interviews and specific learning events. The independent Centre for Development and Environment (CDE) of the University of Berne is conducting this study in December 2008 and results are expected to be available in the first quarter of 2009 – ready to be presented at the 2009 conference on “Perspectives on Impact Evaluation”.
The impact of Fair Trade on cotton producer organisations’ efficiency

Presenter: Gaëlle Balineau

A brief description of Fair Trade. Fair Trade movement aims at improving the lot of small producers in developing countries through the payment of a “fair price” for their products.

FLO1 is the leading standard setting and certification organization for labelled Fair Trade. To be certified and gain access to these markets, producers and traders have to respect some standards (small producers have to be organized in cooperatives which they own and govern democratically; environmental standards restrict the use of agrochemicals and encourage sustainability, etc.). By the end of 2007, there were 632 certified producer organizations in 58 countries, representing 1.5 million farmers and workers.

Fair Trade cotton in West and Central Africa. Fair Trade cotton in West and Central Africa is a project initiated by DAGRIS2 and Max Havelaar France in 2003. Despite their comparative advantage3, African cotton growers are suffering rising poverty: American and European cotton subsidies, as well as productivity gains realized worldwide slump world cotton prices. An unfavourable euro-dollar exchange rate exacerbates difficulties faced by cotton producers who belong to the CFA franc zone. Fair Trade cotton was to enable as many producers as possible to enjoy Fair Trade conditions: certified producers are paid 234 CFA francs per kilogram of seed cotton, whereas in 2007/08, the average producer price was about 160 CFA francs. More than 400 producers organisations in four countries are concerned (Mali, Burkina-Faso, Cameroon, Senegal) and FLO estimates that 28,000 farmers are currently benefiting from the initiative.

The impact of Fair Trade cotton in Mali. A “rapid assessment” (based on group and individual interviews, key informants, and secondary data) of the impact of Fair Trade in Mali was conducted by the author in 2006. Most of the effects previously identified in Fair Trade literature were confirmed: certified producers receive higher incomes, and claim that Fair Trade greatly improved their living conditions.

However, an important question remained unresolved: one of the objectives of the demanding FLO standards is to strengthen producer organisations, requiring them to organize to more effectively manage their own assets, gain access to credit and market outlets, reinforce their exporting capacities, etc.

What is the impact of Fair Trade on the producer organisations’ technical efficiency? It is difficult to answer this question because of the selection bias: organisations which are certified are probably the most efficient ones. In order to control for this selection bias, the author conducted a large scale survey5 on 200 producers organizations in Mali (102 certified organisations and 98 organizations which constitute the counterfactual) and employs two empirical strategies: the first exploits the staggered entry in Fair Trade and uses panel data on five campaigns to identify the impact of Fair Trade on technical efficiency. The second tries to instrument the Fair Trade participation. Empirical estimates suggest that certified producers organizations are more efficient.
The need for change in the scope of Agriculture Research Evaluation: Lessons from strides made in Southern Africa

*Presenter: Mariam A.T.J. Mapila, Choolwe Haankuku*

A shift in the agricultural research paradigm in Southern Africa has been called for by not only donor fatigue but also end users frustration with the lack of real impact of agricultural research outcomes on the livelihoods of its intended beneficiaries. The paradigm shift is not limited to the research processes (from top-down research, conventional approaches to participatory methodologies such as the Farmer Participatory Research as pioneered by the International Centre for Tropical Agriculture) but includes methodologies for setting research agendas and for collecting and analyzing data. The Integrated Agriculture Research for Development (IAR4D) approach is one such research approach but numerous questions abound about the robustness of methods that are being proposed to evaluate its effectiveness in delivering more benefits to end users in comparison to conventional agriculture research approaches and non-IAR4D approaches. Since IAR4D is multidimensional and multi-layered, the proof of concept of its superiority calls for robust quantitative approaches that require large amounts of data for not only communities where it is being implemented but also for counterfactuals.

Studies underway in Malawi, Zambia, Zimbabwe and Mozambique aim to develop such robust quantitative approaches that take into account and endogenous qualitative data and institutional aspects of end users respectively. The primarily objective of these studies is to analyze the impact of IAR4D on the production, marketing and social outcomes of the intended end users. Specifically the studies aim to evaluate the differences in the impact of IAR4D as an approach on the human and social capital development, generation of income and creation of improved livelihoods and the adoption of improved technologies as compared to non-IAR4D approaches and conventional approaches. The main emphasis of the methodology for conducting these studies is the use of robust rigorous quantitative models/approaches for the proof of concept of IAR4D in order to make a contribution towards the field of evaluation of research outcomes in Africa. This methodology would then be used to evaluate any research methodology including conventional and non-IAR4D approaches. Expected outcomes are that the studies will demonstrate that effectively measuring the added value of IAR4D in contrast to non-IAR4D approaches requires a greater understanding of the institutional environment and process of how research is conducted. This understanding would enable the evaluator to better capture the perceived impacts of IAR4D from the end users themselves and other non-traditional research actors such as agribusinesses, policy makers and the end users themselves who become part and parcel of the research process in IAR4D. As a result of this process it is expected that the results of result outcomes would be more effective and feedback into future research agenda's hence ensuring that technologies that are developed are truly beneficial for the end users. In so doing funds from both public and developmental research aid agencies would have more effectiveness in the lives of its intended beneficiaries.
Evaluation of Adoption and Impact of improved sesame crop varieties in Tanzania

Presenter: Deogratias Lwezaura

The focus of the study is to investigate the level and extent of adoption of the introduced new technologies and their impacts to the beneficiaries and communities in the sorghum-based farming system of southern Tanzania.

Purposive and stratified sampling methods were used. A total of four districts were involved in the study. In total 24 villages were covered for interview with 223 farmers interviewees.

The questionnaire paid attention to current crop and livestock production practices and inputs use, level of productivity, level of farm resources available, storage, processing and utilization of farm products and, socio-economic aspects influencing sorghum-based farming systems production decisions within the farming household. In particular, economics of sesame and groundnuts production, farmer preferences on introduced technologies (improved varieties and pest control measures for sesame and groundnuts), adoption behavior, and constraints to adoption. Other information asked was related to factor and output markets, farmers’ income and expenditure levels.

The descriptive statistics assisted in understanding the characteristics of the households. Multiple regression analyses were conducted to deal with issues of endogeneity and simultaneity of explanatory variables. These included Logit (Logistic) and Tobit regression analyses to identify determinants of the extent and the intensity of the adoption of the improved crop varieties respectively. Six out of nineteen explanatory variables tested were significant in explaining the adoption of the improved crop varieties on sesame production. For the case of sesame, they include intervention village type, farmer collaboration, and education of household head and ownership of livestock. Intervention village type, shares of groundnut and maize land were significant variables in explaining the intensity of adoption of improved sesame varieties.

The impact equation was estimated using the actual observations for adoption of improved varieties, as well as the saved residuals from logit regression or adoption equation. The results showed that the average expected income obtained by adopters was higher than that obtained by non-adopters but not significantly different.

The lessons that can be learnt from the study:

- It becomes difficult to deal with attribution effects when you have various players and various people who need the interventions (in this case agricultural improved technologies). Though, it would be time consuming and non-cost effectiveness, but interviewing all various categories would add value to the evaluation.

- Many farmers even in the intervention villages would require accessing the technologies even if they are not project target. The technologies can spread through farmer-to-farmer contacts. Thus, for the intervention of this nature instituting experimental unit or quasi-experimental unit would be not forthcoming.

- Engaging communities in agricultural intervention would bring a significant effect to beneficiaries as they build confidence when dealing with local people.
Evaluating the Impacts of Integrated Agricultural Research for Development Processes: Designing a Participatory Monitoring and Impact Evaluation of the Sub-Saharan Africa Challenge Programme in East, West and Southern Africa

Presenter: Jemimah Njuki

The Sub-Saharan Africa Challenge Program (SSA CP) was initiated in 2004 following extensive consultations with numerous agricultural stakeholders (researchers, extension and development agents, policy makers, farmers, and the private sector) to diagnose the main reasons behind the underperformance of agricultural research in Africa. These consultations established that, besides inadequate funding, the main impediment to the contribution of African agricultural research to development impact lies in the way the research is organized and conducted. The consultations proposed an alternative approach that aims to appropriately embed agricultural research within a larger system of innovation whereby knowledge from numerous sources (comprising all the various actors and stakeholders) is integrated and effectively put into use. This approach to agricultural research is termed Integrated Agricultural Research for Development (IAR4D) and has been adopted by the SSA CP. The core of the IAR4D structure is an Innovation Platform (IP) which brings multiple stakeholders together to who communicate, cooperate and interact to set priorities, develop concepts and plans to promote agricultural productivity and profitability. The SSA-CP is implemented in the Kano-Katsina-Maradi Pilot Learning Site covering Nigeria and Niger, the Lake Kivu Pilot Learning Site covering Uganda, Rwanda and DRC Congo and the Zimbabwe-Malawi-Mozambique site covering the 3 countries. Each of the 3 Pilot Learning Sites hosts 3 Task Force Projects with a total of 36 innovation platforms.

The challenge of the SSA CP is to test the hypothesis that IAR4D performs better than the conventional agricultural research and development (ARD) approaches currently used in SSA. This involves: (i) Defining IAR4D and deriving principles on how it should be implemented; and (ii) Conducting an experiment to empirically establish whether IAR4D is feasible, delivers more benefits to end users, is more cost effective than conventional ARD approaches, and replicable on a large scale. This paper discusses the process of the participatory development of a monitoring and evaluation and impact assessment methodology and tools to answer the key research questions and to evaluate the delivery of outcomes and impacts of the SSA-CP with the teams of the 9 projects. It details the research questions and the framework developed to collect and analyze data to answer the key research questions. The qualitative and quantitative tools and methods for measuring processes and outcomes at different levels (actors/partners in the process, organizations, households and communities) are presented as well as preliminary results from the use of these tools. The paper discusses the challenges of evaluating the outcomes of a methodology / processes vis-a-vis evaluating the outcomes of a project intervention and some of the innovative tools such as social network analysis, outcome mapping and how they can be used in both a qualitative and quantitative perspective.
Impact Evaluation of Agricultural Research (Research for development): A case Study

Presenter: Mohamed M. El-Fouly

The National Research Centre is implementing a program on micronutrient research in soils and crops since 1976. The research results are used to achieve development objectives.

The program started with identifying the magnitude of the deficiency in soils/crops in most regions of the country. In a second stage remedy measures were tested and implemented on large scale on farmers fields. Since then the research program is dealing with issues related to the concept of balanced and integrated fertilization based on soil and crop analysis under the changing conditions of the Egyptian agriculture and cropping systems.

A Fund was initiated in 1986 to supply farmers with advice and micronutrient fertilizers against payment based on soil and plant analysis. The Fund is still operating till now offering its services to farmers. Later on, 1993 and due to the success of the Fund, it was decided to set up and implement a small unit for the production and development of new micronutrient fertilizers to serve the activities of the Fund and assist the industry to develop new fertilizers formula. The progress from each phase to the following one was based on outcome and impact evaluation. This will be explained and discussed.

There were measurable and significant increases in crop production and farmers income. New fertilizer products were developed and local micronutrient fertilizer industry was initiated. Other indirect impacts could be also monitored. Results indicated that these deficiencies might be reflected on the population. Results gained by another group of medical and nutrition researchers could confirm this hypothesis and besides iron, zinc deficiencies in humans were detected.

The paper will describe the different planning steps with the given indicators for outputs, outcomes and impacts during the subsequent phases of the program. Methods of measuring impacts shall also be discussed.

As this program shall continue facing the changing challenges in providing crops with needed nutrients under different conditions, sustainability of the system research/development/farmer/industry had to be considered. Measures for sustainability and reacting to changing needs of the different target groups are given.
Impact assessment of the Uganda National Agricultural Research Organisation

**Presenter:** Sylvester Dickson Baguma  
**Co-authors:** Ponniah Anandajayasekeram, Cyprian Ebong

This paper presents a methodological approach followed in evaluating impact of agricultural research in Uganda since 1994 and demonstrates the application of a range of different approaches that were used. This study started in December 2008 and is expected to be completed in February 2009. The National Agricultural Research Organization (NARO) was enacted a body corporate by the Statute of 1992. This consolidated all public investments in agricultural research in crops, fisheries, forestry and livestock under one semiautonomous institution. NARO received major funding from World Bank in two phases since 1992 to June 2009 for agricultural technology generation, dissemination and institutional development. Despite significant registered achievements in the above areas, there was a need to evaluate the impact that NARO has made on the targeted beneficiaries. To this effect an impact assessment study was initiated.

The study started by training a core group of 30 people from diverse disciplines in the national agricultural research system (NARS) in impact assessment of investing in agricultural research for development. The need to have a comprehensive impact assessment necessitated considering different types of impact notably production impact, economic impact, socio-cultural impact, environmental impact and Institutional impact. Different methods for assessing impacts were analysed. The methods were; Modified Peer Review, User Surveys, Benefit - Cost Methods, Cost-Effectiveness Analysis, Case Studies, Partial Indicators, Integrated Partial Indicators, Mathematical Programming, Simulation Method and Production Function Approach - Econometric Approach to estimates Rate of Return (ROR).

Criteria were identified and used to select 15 research technologies as national case studies for in-depth impact assessment from an inventory of over 218 technologies 70 strategies generated and disseminated to uptake pathways and targeted beneficiaries. In addition impacts of developed policies, capacity development (infrastructure and human resources) and production methods, tools, techniques and processes were evaluated. For each of these an appropriate impact assessment method was identified and study team put together to collect and analyse relevant data guided by the study instruments and write up a report. The different case study reports were synthesised for a final national impact assessment report. Preliminary findings are discussed and the applications of the different methods are compared in the final paper.
Evaluating Impacts of Agricultural Research: Lessons and Challenges Based on the Experience of the CGIAR

Presenter: Mywish Maredia

Public-sector national and international research that contributes to agricultural development has a long and distinguished record of ex post impact assessment (epIA) (a specialized area of evaluation designed to identify and measure the consequences resulting from a program or project’s earlier interventions). This presentation will focus on the tradition of epIA of agricultural research conducted by the international agricultural research centers (IARCs) funded by the Consultative Group on International Agricultural Research (CGIAR). The mandate of CGIAR encompasses research on biodiversity, crops, fish, forestry, livestock, natural resources, and policy. Research is conducted with the aim to generate new knowledge, technologies, practices and policies to achieve developmental goals related to alleviating poverty, enhancing food security, and promoting environmental sustainability.

One of the motivations for doing epIAs in the CGIAR is to meet donors’ demands for accountability. However, there are several distinctive features that make investment in ‘international agricultural research’ and evaluating the impacts of these investments quite unique and different from impact evaluations of other types of development aid projects. These include (i) long time lags in realizing higher level impacts, (ii) the uncertain nature of the potentially large benefits from international agricultural research, and (iii) the role of other institutional partners, policies and programs in transforming international public goods into national and site-specific benefits.

Thus, traditional approach towards accounting for research impact has been to use cost-benefit analysis framework, where benefits are estimated based on primary and secondary data spanning many time periods and geographic areas, making lots of assumptions about the underlying relationships between model parameters, and then estimating economic rates of return (ROR) on project- or program-specific research investment. Also, impact assessment based on success stories is a common practice in the agricultural research evaluation literature in general and of IARC research in particular. Moreover, impact accounting by charging net benefit streams of project success stories to all costs of the program, the institute, and even the system in which they originate, is increasingly practiced.

One of the shortcomings of this approach is that impacts have been too narrowly defined and measured only in units of economic indicators that can fit in a cost-benefit framework and appear to be based on methodology that does not render itself to providing real time feedback to the research management process with the goal of learning from evaluation. EpIAs of agricultural research and this shortcoming have come under scrutiny in recent years with increasing push from donors for evidence of impacts beyond economic ROR and the growing emphasis in the development community on learning from impact evaluations. In light of these developments and the overall objectives of the conference, this presentation will: 1) Provide the historical context and 40+ years track record of ex post impact assessment as practiced in CGIAR both at the individual center-level and at the system-level through the establishment of what is now called the Standing Panel on Impact Assessment (SPIA); 2) Review the lessons learned from this effort on documenting impacts; and 3) Discuss some of
the challenges in the methodological front as the CGIAR system moves towards a new direction of assessing impacts not only for accountability but also for learning purposes, and as the scope of epIA broadens in terms of purpose and goals, the types of research assessed and the types of impact indicators used.

| Theme 1: Improving development effectiveness: the role of impact evaluation | Session number: [5.1] |
| Theme 2: Impact Evaluation Approaches and Methods | |
| Session stream: Approaches and methods | Session name: Approaches and methods |

**Beyond feedback loops: systems techniques for impact evaluation**

**Presenters: Bob Williams, Irene Guijit, Patricia Rogers**

Systems approaches can be useful in addressing many of the tasks in impact evaluation - negotiating the focus of the evaluation, conceptualising the intervention and how it and other factors might contribute to the impacts, describing, measuring and valuing actual impacts, reporting and supporting use of findings. In particular they are useful in helping evaluators to ask, and answer, better questions – and questions are at the heart of impact evaluation. However many people are only familiar with one of these techniques - causal loop diagrams. This session provides an overview of a range of systems approaches, including soft systems critical systems heuristics, and complex adaptive systems, and how these can be applied to address evaluation questions that are highly relevant to impact evaluation in developing countries.

| Themes 2: Impact Evaluation Approaches and Methods | Session number: [5.2] |
| Session stream: Approaches and methods | Session name: Approaches and methods |

**Reporting on outcomes together: The performance story reporting technique**

**Presenter: Jessica Dart**

Decisions makers are often under pressure to generate relevant and useful reports concerning the extent to which a program has contributed to outcomes within tight timeframes. The innovative new technique named ‘Performance Story Reporting’ (PSR) aims to address this need. Building on the seminal work of John Mayne’s ‘contribution analysis’ PSR presents a framework for reporting on a program’s contribution to long-term outcomes using mixed methods and participatory process. The process steps include clarifying the underpinning intervention logic, developing guiding questions for a social inquiry process and data trawl. Final conclusions about the extent to which a program has contributed to outcomes are made by an ‘outcomes panel’ and recommendations are developed at a large group evaluation summit workshop. In addition to the process, PSR offers a five-part performance report template.

This technique has been successfully used with 14 studies in Australia (In all States and Territories as well as with Indigenous sea and land management programs in the Northern Territories) and has become a favoured tool for reporting on outcomes for several Australian Government Departments. It has recently been used in a variety of more diverse settings including, Uganda, New Zealand and PNG.
Evaluation in the Practice of Development

**Presenter: Martin Ravallion**

While standard methods of impact evaluation can be useful, they often leave significant gaps between what we know and what we want to know about development effectiveness — gaps that stem from distortions in the market for knowledge. The paper discusses how evaluations might better address these knowledge gaps and so be more relevant to the needs of practitioners. It is argued that more attention needs to be given to identifying policy-relevant questions (including the case for intervention); that a broader approach should be taken to the problems of internal validity; and that the problems of external validity (including scaling up) merit more attention by researchers.

Anyone who doubts the potential benefits to development practitioners from evaluation should study China’s experience at economic reform. In 1978, The Communist Party’s 11th Congress broke with its ideology-based view of policy making in favor of a more pragmatic approach, which Deng Xiaoping famously dubbed the process of “feeling our way across the river. At its core was the idea that public action should be based on evaluations of experiences with different policies: this is essentially what was described at the time as “the intellectual approach of seeking truth from facts” (Du Runsheng, 2006, p.2). In looking for facts, a high weight was put on demonstrable success in actual policy experiments on the ground. The evidence from local experiments in alternatives to collectivized farming was eventually instrumental in persuading even the old guard of the Party’s leadership that rural reforms could deliver higher food output. But the evidence had to be credible. A new research group did field work studying local experiments — non-randomized experiments — on the de-collectivization of farming using contracts with individual farmers. This helped to convince skeptical policy makers (many still imbued in Maoist ideology) of the merits of scaling up the local initiatives (Xiaopeng Luo, 2007). The rural reforms that were then implemented nationally helped achieve probably the most dramatic reduction in the extent of poverty the world has yet seen.

Unfortunately we still have a long way to go before we will be able to say that this story from China is typical of development policy making elsewhere. Practitioners — working in developing countries, donor countries and international agencies — search continually for operational solutions to pressing development problems. Researchers have the training and skills needed to prove the tools for informing that search, and help learn from our successes and failures along the way, as recorded knowledge accumulates. However, practitioners rarely appreciate the full benefits from rigorous evaluations. And, as a rule, it is the development practitioners who hold the purse strings. The result is that too little rigorous research of relevance to development gets done. We know too little about the efficacy of development efforts and the learning process becomes too weak to reliably guide practice. The outcome is almost certainly less overall impact on poverty.

This paper argues that too little rigorous impact evaluation on development effectiveness gets done and that the evaluations that are done currently are not as useful as they could be for learning about development effectiveness. The production of evaluations is riddled with problems of externalities, selection-biases in funding, myopia, publication biases and weak incentives to innovate in developing and adapting the tools needed. Distortions in the “market for knowledge” about development effectiveness leave persistent gaps between what we know and what we want to know. The paper points to a number of things that need to change if the potential for evaluative research to inform development practice is to be fulfilled.
Alternative approach to monitoring and evaluation based on a case study from Somalia

**Presenter: Demetrio Innocenti**

The M&E/MIS system has been developed using open source technologies, for content management, surveys and e-forms, document management, and project monitoring.

The key concept is how this system can be replicated and mainstreamed within national Ministries and Agencies, particularly of Developing Countries, to provide an efficient and highly cost-effective MIS system that can allow tracking of data and information for impact evaluation of rural development interventions.

Key technical solutions that will be presented:

**CDRD Portal** is a web-based dynamic portal engine and content management system. The system enables non-technical users to manage content on the website. Using a simple, browser-based interface the users will be able to easily publish news, press releases, files, images, and create an unlimited amount of sections or content pages on the CDRD site. The portal is available online at www.somcdrd.org and can be used as a platform to promote collaboration among the different partners.

**CDRD Project Management System (PMS)** is a web-based application designed to assist in project scheduling and monitoring. The system includes modules for communities, sub-projects, tasks (with Gantt charts and can be exported to MS-Project), committees, Galley, calendar, contacts and multi-language support. The PMS incorporates a Document Management System that enables users to upload/download files (i.e. Excel, Word, pdf, PowerPoint, etc) and has a modular design that allows extra modules (such as procurement and inventory) to be added in.

The structure of the PMS is finalized to multiple purposes:

- To keep track of daily monitoring tasks at field level such as GANTT charts for resource mobilization and tracking, inputs delivery and outputs achieved. (See attached screenshots)
- To document and report progresses against resource employed using a series of results based indicators. All this is functional to have an easy snapshot of the project at any given moment: which results have been archived at what costs.

**CDRD E-Forms & Survey Management System (EF&S)** is a web-based application that enables our social development experts and field facilitators -with no coding knowledge- to develop, publish and collect responses to multi-lingual e-forms or surveys. The system supports many different question types, ranging from the simple Yes or No variety to choices such as 10-point scales, open-ended answers, or dropdown choices. There is no limit to the number of questions, and the system also offers ways to automatically email, notify, invite or remind users to fill out an e-form/survey.

The forms can be directly filled by the facilitators in the moment in which they interview key informants and then synchronizing through internet. At real time the data can be automatically analyzed through predefined statistical tools (i.e. frequency analysis) built in EF&S or exported to SPSS, text, CSV, or Excel files. EF&S is available online through the CDRD portal and can support the streamlining and automating of data-collection of different processes such as project proposal forms and M&E indicators.
Application of Logic Modeling to Evaluate Impact

*Presenter: Kafula Silumbe*

*Co-authors: Ralph Renger*

This paper presentation will provide a framework for understanding and developing logic models using the ATM approach. It will include an overview of the three step logic modeling process which has been implemented throughout the United States, Canada, and Africa and show its utility in program intervention and policy. The focus of the presentation is to allow participants an opportunity to learn the skills necessary to develop a logic map for a problem of interest (step1), prioritize the underlying conditions appearing in the logic map for strategy development (step 2), and identify potential measures to assess the underlying conditions (step 3). Participants will understand how to use the outcomes of each step of the process to quickly complete the traditional logic model table often required of programs by funding agencies. Overall, utilization of the logic modeling process will help to ensure programs have the best chance of producing its intended outcomes.

Poverty analysis on the ground: Challenges of developing a monitoring framework for Impact Evaluation (case of community dynamics in Democratic Republic of Congo)

*Presenter: Mashika Bin Kasongo Albert*

Vers la fin des années 1999, la BM et le FMI optent pour une révision de leurs approches d’intervention dans les pays en voie de développement. Ils leur proposent désormais de se doter de DSRP/CSLP élaborés de manière participative avant d’accéder aux facilités telles que la remise de leur dette (Initiative Pays Pauvres Très Endettés), Facilités pour la croissance et la réduction de la pauvreté.

Pour rompre avec les méthodes de planification dirigiste et bureaucratique, la RDC décide, en 2003, de rendre la parole à la population congolaise pour analyser sa propre situation et recueillir des données qui allaient lui permettre de se doter d’un DSCRP de qualité, bénéficiant en plus d’une large légitimité. C’était là l’objectif des consultations participatives sur la pauvreté auprès des communautés de base (Analyse Participative de la Pauvreté ou APP) dans 94 territoires et communes de la RDC. L’exercice a été confié sur une base compétitive à 10 organisations de la Société civile dont la CDC (20 territoires dans 2 provinces). Le processus a touché au total presque 35.000 personnes.

La collecte des données s’est faite en recourant aux outils et techniques de la Méthode Accélérée de Recherche Participative (MARP); cela a permis aux populations locales de (1) analyser, partager et améliorer leurs connaissances sur leurs conditions de vie et (2) déterminer et programmer leurs priorités dans des plans d’action communautaires en vue d’agir, de suivre et d’évaluer plus tard les réalisations.

La démarche est passée par les étapes suivantes : (1) Formation des équipes au niveau national et ensuite provincial, (2) Identification des villages et quartiers dont les habitants devaient être consultés, (3) Descente des équipes MARP sur le terrain : en six jours, dans chaque site (village ou
les équipes ont organisé des Focus Groups. Certains leaders d’opinions tels que des chefs de villages ou pasteurs des églises, ont été interviewés individuellement. Par la suite, un groupe hétérogène de 10 habitants (femmes et hommes) du village/quartier a été constitué pour partager avec l’Equipe MARP la synthèse des informations recueillies en vue de leur validation et appropriation par les concernés. Les conclusions retenues ont été consignées dans un Procès-Verbal contresigné par tous les participants. Des ateliers de validation ont été organisés au niveau territorial/communal (94), provincial (11) et national (01).

Le Ministère du Plan a mis sur pieds 15 Groupes thématiques (cadres de concertation pour la mise en œuvre de la stratégie à travers un Programme d’Actions du Gouvernement) dans lesquelles les organisations de la Société Civile (OSC) participent. Les 10 OSC ayant conduit l’APP se sont regroupées en une plate-forme dénommée NADYC (Noyau d’Accompagnement de la Dynamique Communautaire) dont l’objectif est de faire un suivi indépendant de la mise en œuvre de la stratégie de réduction de la pauvreté.

L’expérience du NADYC en la matière lui a permis d’identifier les défis dans les domaines suivants: (i) l’échelle de participation et du vrai débat sur l’impact; (ii) l’instrumentalisation des communautés; (iii) l’efficacité; et (iv) l’articulation entre la gouvernance locale, la décentralisation et le développement participatif.

### Themes 2: Impact Evaluation Approaches and Methods

Unpacking the notion of program theory from indigenous and cross-cultural perspectives

**Presenter:** Sanjeev Sridharan

This presentation is an international collaboration between researchers based in Canada, South Africa and Japan.

The growth of methods to integrate program theories with statistical designs in impact evaluation is partly based on the old statistical adage "no causation without a theory". This presentation will attempt to unpack what it means for programs to work from different cultural lenses and how such a cultural understanding of key evaluation concepts can help design more influential impact evaluations. Our argument is that while program theoretic concepts may be stable across cultures, different cultures groups frame and represent program theoretic concepts differently. As example, it is possible theories of change/programme theories might need to be framed and represented differently for an intervention focussed in an aboriginal group in Canada as compared to a similar intervention in a rich area in Nairobi.

Additionally, it is our belief that most academic concepts in standard program theories are quite sterile – there is a need for practice based thinking (from mainstream, indigenous, and international settings) to inject richer vocabulary on how we conceptualise the pathways by which programs work. This presentation will be based on a project that attempts to build an international dialogue on practice based insights on program theory.

The key concepts that will be covered in this presentation include: (i) Program theory/theories of change; (ii) Context; (iii) Timeline of impact; and (iv) Relationship of program theories to sustainability.
Making Evaluation of Impacts more Robust using Contribution Analysis: the Case of European Development Aid Support

**Presenter: Eric Monnier**

Contribution analysis investigates the chain of assumptions which connects the evaluated public intervention and its intended impact. Data are gathered by any possible means (inductive as well as deductive) in order to develop a step-by-step chain of arguments asserting that the policy has (or has not) made a contribution, and possibly ranking the programme among other contributions. The reasoning is then submitted to systematic criticism until it is carefully confirmed. It is then presented in the form of a short text and a diagram.

We have been developing this analysis strategy for several years, for the benefits it provides in terms of: feasibility in nearly all evaluation contexts, compatibility with all types of evidence quantitative and qualitative, capacity to analyse specific impacts in a conclusive way, possibility to be subjected to a rigorous quality assessment.

Contribution analysis is one of the approaches through which evaluators can reach robust conclusions on cause-and-effect issues. It may be used in a wide range of assignments, contrary to other approaches for impact analysis so called “attribution analysis”, as randomized control trials, or econometric modelling.

The practice of contribution analysis is however still limited in quantity as well as in quality. Thus there are not yet developed professional standards and there is a risk that the term becomes just an excuse for not doing rigorous impact evaluation notably in the context of complex development aid programmes.

Development aid is increasingly organized through sector wide programs and general budget support. These new approaches depart from the traditional accumulation of unconnected projects. For the supervisory or accountability functions, such as evaluation, the change goes the other way round: from the simple to the complex. The logical framework tool tends to give each project a clear purpose, a well identified target group, a clearly set of activities logically organised. On the contrary, the evaluability of large scale development aid programmes is questionable because they involve multiple activities, multiple target groups, and an intricate web of cause-and-effect assumptions.

The difficulty of assessing such programmes is discussed on the basis of the evaluation of the overall European Commission support to Tanzania (EC, 2006). The speech will present several approaches that cope with the “evaluability barrier”.

The Impact of Evaluation Outcome Measurement as a Feedback Mechanism for Sustainable Development

**Presenter: Apollos B. Goyol**

The dividends of democracy are evident when the electorates can clearly identify and feel the impact of the resources of the government actually produce tangible benefits for the people.
That is to say, they demand better accountability for the use of resources and a feedback mechanism for sustainable development.

In addition, since accountability and transparency is the order of the day, the main objective of this paper is to aid the Federal, State, and local government officials to understand the essence of accountability. This is best carried out by using the Measuring Program Outcomes Mechanism, relevant to their various agencies and ministries. I believe this is in line with the agendas of the various levels of government as promised to the electorates.

Once clear and compelling answer to the question of “Why measure outcomes?”, is to see if programs really make a sustainable difference in the lives of people. However, although improved accountability has been a major force behind the move to Outcome Measurement, there is an even more important reason: to help programs improve services. Outcome Measurement provides a learning loop that informs programs on their effectiveness. This offers findings we can use to adapt and improve, thereby becoming more efficient and effective which ultimately results in sustainable development.

<table>
<thead>
<tr>
<th>Themes 2: Impact Evaluation Approaches and Methods</th>
<th>Session number: [5.6]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session stream: Approaches and methods</td>
<td>Session name: Approaches and methods</td>
</tr>
</tbody>
</table>


**Presenter: Jos Vaessens**

The reasons for doing impact evaluations are many but the ultimate reason for promoting impact evaluations is to learn about 'what works and what doesn't and why' and thus contribute to the effectiveness of development interventions. The purpose of NONIE is to promote more and better impact evaluations among its members. The Membership of NONIE is of the opinion that better coordination, sharing of methods, and mutual learning-by-doing would be beneficial to the promotion and practice of IE. The current Guidance document was developed to help provide global standards and direction on high quality and effective impact evaluations. The Guidance seeks to provide clarity on potential areas of misconception in impact evaluation and to offer guidelines for alternative methods and approaches for ensuring the rigor of evidence and raising IE quality.

The Guidance argues that there is no universally accepted definition of rigorous impact evaluation and that rigor is more than just about methodological design. Rigorous impact evaluation requires addressing the key issues in impact evaluation in an appropriate manner. The Guidance document is structured around nine key issues in impact evaluation (described below). The first part of the document, comprising the first six issues deals with methodological and conceptual issues in IE and constitutes the core of the Guidance document. In addition, the second part focuses on managing IE and addresses aspects of evaluability, benefits and costs of IE and planning.

Nine key issues in impact evaluation:

1. Identify the (type and scope of the) intervention
2. Agree on the objectives of the intervention that are valued
3. Carefully articulate the theories linking interventions to outcomes
4. Address the attribution problem
5. Build on existing knowledge relevant to the impact of interventions
6. Use a mixed methods approach: the logic of the comparative advantages of methods
7. Determine if an IE is feasible and worth the cost
8. Start early – getting the data
9. Front-end planning is important

The presentation will lay out the main messages of the Guidance on these nine most important issues in impact evaluation.

**Themes 2: Impact Evaluation Approaches and Methods**

**Session number: [5.7]**

**Session stream:** Approaches and methods

**Session name:** Approaches and methods

**Doing Impact Evaluations with Ex-post or Independent Data: Challenges for Evaluators**

**Presenters: Ximena Del Carpio, Dominique Van De Walle, Kathleen Beegle**

Evaluators face several challenges when trying to measure the impact of a program. Among the crucial challenges is having adequate data to be able to have a valid counterfactual and measure the appropriate set of outcome indicators. In an optimal scenario, a program collects a baseline before the program starts and one or more follow-up rounds, using a survey instrument designed around the program objectives. The data cover both treated and non-treated areas and the timing of the post-intervention surveys allow sufficient time to capture full program impacts but avoids contamination from spillover effects. The challenge of not having purposefully collected data and having to rely on data collected ex-post or independently of the program and worries about timing are salient among evaluators. Researchers have developed techniques that allow evaluators to face these challenges and makeup for less than ideal data; moreover, a growing number of data-sets (e.g. regional, national) allow evaluators to apply these methods and measure impact of various development programs. This session is intended to motivate practitioners to address the concern of not having purposefully collected data by presenting recent examples from around the world where alternative sources of data or data collected ex-post(124,191),(658,303) coupled with sophisticated methods, was used to successfully evaluate development work. The session also presents some evidence on how well these techniques and short cut methods work in practice.

The session will include three impact evaluation presentations. Presenters are well versed in conducting impact evaluations under challenging circumstances, particularly focusing on the use of independent data sources or data collected after the program started. Papers are geographically and thematically diverse. Data sources explored will vary; examples are: national household surveys, ex-post project data (collected after a project started), geographic information and qualitative data. Methods of approach in this session range from propensity score matching, nearest neighbor matching, double-differencing, pipeline approach, and mixed-methods.

**Themes 2: Impact Evaluation Approaches and Methods**

**Session number: [5.8]**

**Session stream:** Approaches and methods

**Session name:** Approaches and methods

**Mixed methods in impact evaluation**

**Presenters: Jennifer Greene and Mel Mark**

How should development evaluation practitioners make practical and defensible judgments about what constitutes worthy evidence on program quality and effectiveness? Which impact evaluation designs and methods can best generate credible evidence, under what conditions?
In a follow-up to the plenary on credible evidence and method, this session will invite participants to work through practical design decisions and to share their resulting ideas. The session will feature a diversity of development contexts and questions for discussion.

Jennifer Greene and Mel Mark will facilitate discussion in small groups, guide the sharing of ideas, and discuss design alternatives for the impact evaluation contexts considered during the session.

**Themes 2: Impact Evaluation Approaches and Methods**

**Session number:** [5.9 ]

**Session stream:** Approaches and methods

**Session name:** Approaches and methods

**Agden Roundtable on Impact evaluation: Approaches and methods**

**Presenters:** Adeline Sibanda, Florence Etta, Akosua A. Ampofo, Nestorine Campaore, Oloronke Ladipo, Sarah Okwaare

The African Evaluation Association (AfrEA), the Network of Networks on Impact Evaluation (NONIE) and the International Initiative for Impact Evaluation (3ie) have joined forces to host a conference whose theme is PERSPECTIVES ON IMPACT EVALUATION: Approaches to Assessing Development Effectiveness.

The Africa Gender and Development Evaluators Network (AGDEN) is submitting a proposal to host a round table on: Impact evaluation: Approaches and methods.

**Background.** In 2006, as a direct result of recommendations at a number of meetings and consultations on the New Aid Modalities as they have become popularly known, UNIFEM established an expert group to interrogate this new terrain and suggest how gender equality would best be achieved within the new landscape dominated by the Paris Declaration (PD) of March 2005. The Africa Gender and Development Evaluators Network (AGDEN) was invited to join this group to contribute specifically in the area of developing gender sensitive indicators and tools for use in the monitoring and evaluation the new aid mechanisms.

The expert group on the new aid modalities created by UNIFEM in 2006 produced a set of gender sensitive indicators together with an accountability and assessment framework for gender equality of the Paris Declaration and other aid instruments. These indicators have been reviewed and revised by AGDEN.

The PD is currently the guiding framework for Aid Effectiveness. To be effective, aid in all its modalities must of necessity support gender equality as no aid can be said to be truly effective if it does not bring about gender equality, the advancement of women and the protection of human rights. While many of the major reviews of the PD to date suggest that the narrow focus on efficiency and structural reforms of aid mechanisms has limited positive impact on development effectiveness in general and gender equality in particular (DCD/DAC, 6-2006, 7), others see the opportunities offered by the gaps as valuable entry points for the work of engendering the commitments, indicators as well as targets (DCD/DAC, 2007, Fleming S. et al, 2007) on one hand and aid effectiveness in general.

**Aim.** The aim of the Roundtable is to share the gender responsive indicators and discuss the use of these indicators in measuring impact of the Paris Declaration and aid effectiveness in general. It is hoped that more evaluation practitioners will take the indicators on board. This discussion will also set a platform for Paris Declaration Evaluation Phase II which will be conducted by the Reference Group starting 2009.

**Methodology.** One presentation will be made in this roundtable, followed by four discussants. All the participants are members of AGDEN.
Using randomized evaluation to test development effectiveness

Presenters: Annie Duflo, Rebecca Thornton, Kamila Gumede

Randomized evaluations are considered the “gold standard” in program/impact evaluation. Until recently they have been used relatively infrequently in the evaluation of poverty programs. This is partly due to the perceived difficulty in introducing an element of randomization into the implementation of poverty programs in a way that is both acceptable to recipients of the program and logistically feasible for implementers. However, recent innovations have demonstrated how the rigorous scientific demands of randomized evaluations can be made compatible with the logistical and practical constraints of implementers working on the ground.

J-PAL and IPA have acquired considerable expertise in running randomized evaluations: the professors in our network have developed many of the techniques that have allowed the expansion in the use of randomized evaluations of poverty programs. We have 181 completed or ongoing randomized evaluations in 30 countries on a wide range of subjects and have trained 441 practitioners through our executive education courses.

The propose panel brings together J-PAL partners and researchers to present a wealth of experiences in conducting randomized evaluations. It will provide participants with an understanding of the practicalities of conducting randomized evaluations and the wide range of policy issues that can be evaluated using this rigorous methodology. J-PAL experts, affiliates, and partners will share their experiences from different sectors including: education, microfinance, reproductive health, community driven development, labor market reforms, governance reforms and corruption. The objective is to illustrate (through practical examples) that the boundaries of intractable questions are rapidly being pushed back and that randomized evaluations are more feasible and adaptable than many realize.

Enjeux et défis de l’évaluation d’impact en Afrique : Cas d’une évaluation réalisée au Bénin

Presenter: Serge Djossa Adoun

Co-author: Marie Gervais

L’évaluation est aujourd’hui un outil de gestion et de prise de décision incontournable dans les organisations. Mais si cette conviction de la valeur de l’évaluation est partagée sur l’échiquier international, dans bien des pays, tout est encore à faire en matière de culture d’évaluation. Ainsi, en Afrique, nombreux sont les politiques, programmes et projets dans lesquels les gouvernements investissent chaque année, mais dont on n’a jamais évalué les effets, les impacts. Aussi, nombre d’organisations, dans les pays africains en l’occurrence, ne disposent d’aucun système de suivi et d’évaluation pour mesurer les performances, soutenir l’amélioration des services, engendrer un changement dans le milieu organisationnel et contribuer plus efficacement au développement.
Comment l’évaluation d’impact doit-elle être conduite dans un tel environnement? Quels indicateurs utiliser? Quelles approches méthodologiques privilégier? Comment intéresser et mobiliser les bailleurs de fonds et autres groupes concernés par l’évaluation? Comment alors travailler au développement et à la pérennité d’une culture d’évaluation?

Cette présentation retracera les grandes lignes d’une démarche d’évaluation réalisée sur un programme de santé sexuelle et reproductrice des adolescents et jeunes au Bénin. Les caractéristiques de l’approche d’évaluation et la méthodologie associée seront abordées. Les défis et les enjeux rencontrés tant dans la préparation du terrain que dans la collecte de données seront discutés. Un aperçu des résultats sera également présenté. Finalement, une discussion suivra concernant les leçons dégagées de cette expérience conduite dans un pays où les acquis en évaluation sont encore fragiles.

**Themes 2: Impact Evaluation Approaches and Methods**

**Session number: [5.11]**

**Session stream: Approaches and methods**

**Session name: Approaches and methods**

---

**Applying Quality Standards in Impact Evaluation: Case of CARE Program Quality Framework and Evaluation Policy**

**Presenter: Ahmed Ag Aboubacrine**

This panel discussion intends to present the success and challenges in the use of CARE International Program Quality Framework (Vision/Mission, Core Values, Programming Principles, Design, Monitoring and Evaluation Standards, Core Guidelines; Technical Guidelines). This will includes first a presentation of the CARE standards and their relation to the ones of OECD-DAC and also the combination of those standards with specific donors’ requirements and donor-based constraints.

Secondly the panel will explain the evolution of thinking within CARE about standards since they have been adopted in 2002 (Jim Rugh 2004). The lessons learnt in the utilization of CARE standards are leading to a significant shift (from Project to Programme approach) in the way CARE is working.

Thirdly, the panel will show how CARE staff uses the standards in donor quality assurance and compliance. That will include successes and failures as well as the challenges encountered during the use of the standards from the design of an intervention to its final evaluation.

Some highlights will be about constraints related to understanding the standards, time, costs and staff capacity. An emphasis will be putted on what is called here “human factor” which influence a lot the quality of utilization of DME standards and how that affect the achievement of impact. The human factor constitutes a bias that comes from various directions and can have many dimensions depending on the agenda, skills and competencies of the key stakeholders involved in designing, monitoring and evaluating development intervention.

A special focus will be presented on the institutionalization of those standards in post-conflict countries such as Sierra Leone where the utilization of evaluation to influence decision makers is far from being a reality despite the existing potential (Paris Declaration, Accra Agenda for Aid Effectiveness). That emerging trend of promoting a culture of evaluation and accountability is also impacting the way aid agencies work and use their quality standards.

Finally, the panel will discuss some a list of “Do” and “Don’t” in applying quality standards. Those points may be applicable to all evaluation standards and guidelines in development interventions.
Feminist Research Perspective on Impact Evaluation

**Presenter: Helen Kezie-Nwoha**

Debates about research relying on the experiences of women have dominated feminists’ discourses for years. Feminists have challenged traditional research methods that ignore women’s experiences, and argue that for any research to lead to improved lives of women it must privilege women’s voices. Thus giving rise to feminist ‘epistemologies’, indicating its diversity; this also has raised interesting debates amongst feminist researchers. The challenge for feminist researchers has been how women’s knowledge can be produced without biases and not privileging the researcher as the ‘knower of knowledge’. In terms of evaluation, privileging women’s voices becomes a strategic alternative, as it is a reflection of the real impact of project/programme on women’s lives.

In assessing the impact of a community gender training programme that challenged cultural and traditional practises which promotes gender inequality, my experience is that women’s perspectives are important in determining the impact of projects/programme and if any changes have taken place.

In adopting a feminist perspective, interviews and observations are the common research methods applied. I find interviewing engaging as it offers access to people’s ideas, thoughts, and memories in their own words, this is particularly important for the study of women, because it solves the problem of centuries of ignoring women in research. Women are also ignored in impact evaluation processes. For impact evaluation to be truly evident it must get the perspectives of both women and men, in this respect data collection has to be representative of all categories of project beneficiaries.

A participatory evaluation process which engages all categories of people adopts qualitative methodology, where researchers study spoken representations; it also involves recording human experiences as a means of data collection (Punch, 2005: 168). The challenge with interviews is the tendency to present assumed perspectives. This has the tendency to affect data analysis. In discussing data analysis Punch points out that researchers are faced with the challenge of representation, this is reflected in the written data, which has the potential to reflect the researchers constructed ideas (2005:58). To what extent women’s experiences can be a reliable source of knowledge in any research has been a complex and debated issue among feminist researchers. During such interviews there are concerns over the validity of the different versions of accounts of women’s lives from the women I interviewed. My position on the issue is that women are in a position to relate their experiences and give account of how projects have impacted their lives. Such experiences are useful for project assessment and planning. Evaluators therefore have to recognise the importance of listening to women voices and the voices of other marginalised groups during impact assessment to ensure that evaluations measure the benefits of programmes.
A roundtable discussion on improving effectiveness of public service delivery through impact evaluation

Presenter: Paul O Omondi

Background. In 2006, the Kenya government established The New Performance Appraisal System (PAS). Notably, since inception, PAS has expanded from a pilot group of 16 commercial public enterprises in 2004, to eventually cover the entire public service in Kenya. As part of an integrated performance evaluation process, PAS is based on the recognition of the need for efficient, effective and quality public service delivery.

Citizens Service Delivery Charter in the Context of Impact Evaluation. PAS requires every public institution to outline its Citizens Service Delivery Charter, which empowers the public to expect and to demand specific service standards. As such, the institutions are accountable to the public, not just for what has been done but more so for whether or not they have accomplished their goal in line with their charter.

The Citizens Service Delivery Charter contains three key elements; mission and vision statements and long-term strategic goals. The PAS measures these components based on the following parameters: Performance & quality of service; Involvement of service consumers; Evaluation of quality of service delivery; and Ranking of institutions by excellence in performance by using composite score.

Impact Evaluation Methodologies. Central to performance evaluation is application of customer satisfaction surveys in order to establish levels of satisfaction with Government Services. Policymakers have recognized customer satisfaction surveys as important tools for monitoring and reporting of program accomplishments.

In Kenya, therefore, public institutions are required to carry out regular customer satisfaction surveys to ensure that the quality of customer service improves progressively. Each year should demonstrate growth in the performance of each indicator over the previous year. The Government does not evaluate its own performance, rather evaluation of the Ministries and State Corporations are carried out by experts drawn from the private sector, professional associations, the academia and top business executives. These surveys are directed to segments of the society who have interacted with the institution. As aforementioned, other than simply providing results to show the level of compliance with service standards, performance evaluation in this regard allow the institutions to monitor their performance over time, results of which are used to inform further action for improvement where necessary.

Use of Impact Evaluation in Improving Public Service Delivery. The Government of Kenya has increasingly recognized the need for evidence based policy. In examining the relationship between evaluation results and policy making, we take note that the findings generated by surveys are used to come to more informed decisions about the likely impact of policies in Ministries and state corporations.

The government takes cognizance of the fact that, in order to achieve sustainable service delivery, the main focus of the Government is to ensure full integration and comprehensive cascading of the Performance Contracting Process in all Government Institutions. Granted, insufficient capacity among some government officials and policy maker to interpret and utilize research remains a challenge. This suggests the need, first, to develop capacity in terms of educating officials about evaluation and second, to provide research findings in easy to understand manner so as to ensure they are beneficial.
The Design, Implementation and Use of Rigorous Impact Evaluation

Presenter: Paul Gertler

Experiences in Impact Evaluation: the PEMA Perspective

Presenter: Laila Shahd

As a background, PEMA has over six years experience as national Impact Evaluation agency in Egypt. Projects evaluations have been selected from several sectors and development themes including: education, public health, industry, agriculture, environment, water supply, as well as “girls education”, “eco tourism”, and vocational training.

The proposed presentation would highlight PEMA’s experience in terms of the evaluation methodology used, main findings, and recommendations.

In this respect, the presentation would also provide a brief account of how joint evaluation initiatives were carried out with CIDA, Italian Cooperation, and other funding agencies.

The presentation would conclude with a summary of the key lessons learned.

Impact evaluation in the light of Action Research: methodological issues and perspectives

Presenter: Ahmed Bencheikh

Impact evaluations of development activities, particularly social policy in Morocco (fight against poverty, empowerment of women, violence against women and children, etc.) have become crucial in Morocco today. Generally, evaluations, although qualified as impact evaluations deal mostly with outputs. At best, they are interested in the primary effects of programs and projects.

Impact evaluation studies are most often confronted with a methodological challenge, which characterizes the scope of changes on the beneficiary populations, territories and societies concerned. In addition, social work is faced with the consequences of competitive strategies developed by different institutional or international intervenes, sometimes on the same fields...

Based on theoretical and empirical contributions of Action Research our objective in this paper is to examine the impact evaluation studies in social and human development in Morocco. The idea is to question these assessments in the light of the main action research analyzers, namely:
(i) First, the knowledge of the environment in terms of quality of background studies (Baseline study, diagnosis, feasibility study, etc.);

(ii) Furthermore, the process of intervention on the social and territorial ground reflected mostly by the mobilization approaches of beneficiaries and partners of development actions;

(iii) Finally, the strategies developed by the State, local NGOs, national and international stakeholders (donors, local authorities, etc.).

(iv) The analysis we propose aims the understanding of the impact evaluation process in the light of analytical assumptions of Action Research:

(v) To what extent Action Research can improve the methods and approaches of impact evaluation?

(vi) What complementarities can be established between impact evaluation and Action Research?

(vii) How Action Research can make impact evaluation more rigorous?

<table>
<thead>
<tr>
<th>Themes 2: Impact Evaluation Approaches and Methods</th>
<th>Session number: [5.15 ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session stream: Approaches to M&amp;E: MENA experiences</td>
<td>Session name: Approaches to M&amp;E: MENA experiences</td>
</tr>
</tbody>
</table>

Measurement of Process Effectiveness – the Use of Dynamic Indicators

**Presenter: Meera Sundararajan**

Indicator development is a core activity in building any M&E system. An indicator is a specific measure that when tracked systematically over time, indicates progress (or not) toward a specific target. However design of indicators for tracking change towards a dynamic target, poses a challenge, in interventions like literacy projects which are process driven resulting in human capacity oriented changes.

**The context.** CARE India’s tsunami response program in southern India, initiated an adult literacy program across 40 villages covering 30 adult female learners in each centre. The core objectives of the project are:

- To impart and improve functional literacy and numeric skills
- To provide learning experiences which promote better livelihood preparedness

The course period for a batch of 1200 learners consists of a core learning period of 6 months and is divided into 3 phases each focusing on the following inputs:

1. Phase-1: Basic Literacy & Numeric Skills with Projects/Practical
2. Phase-2: Reading, Writing & Calculation Skills with Projects
3. Phase-3: More of Projects for 'Experiential Learning'

Each phase has eight weekly teaching guides (except the last one which has 2 weekly guides) with each guide being focused on the learning objectives for the week.

**The monitoring framework.** While traditionally monitoring of literacy programs are focused towards class attendance, this initiative sought to go beyond and look at the learning process. Learning processes, however are dynamic with the learner moving from one learning milestone to the next during the course of achievement of learning objectives.

Therefore each phase focused on specific learning indicators which are in line with the phase learning objectives. However the project also recognizes the final outcome from the 6 month learning program as a result of which there are final outcome level indicators -milestones which are completely assessed throughout the third phase –at the end of which the overall learning objectives are to be met. A baseline is generated on the final learning milestones at
the time of initiation and assessed again at the end of phase three. The initial two phases however assess the process indicators alone.

Data on process indicators are reported on a fortnightly basis where the teacher assesses each learner through the following methods:

A. Written Testing
B. Classroom observation
C. One to one interaction

Based on the assessment, learners are rated into 4 categories depending on the extent to which each milestone is being met by individual learners (absolute level of measurement) as well as how each learner fares with regard to the best abilities expressed at that learning centre (relative measurement). Centre rating is done based on the number of learners who are in the outstanding to normal levels at learning milestone achievement. Centers which exhibit high numbers coming in the category of being unable to meet milestones, are reviewed with the teachers using supporting attendance data to understand reasons for poor performance. Additionally, data on socio economic profiles are used to understand reasons of poor performance despite regular attendance and participation in class work.

<table>
<thead>
<tr>
<th>Themes 2: Impact Evaluation Approaches and Methods</th>
<th>Session number: [5.15 ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session stream: Approaches to M&amp;E: MENA experiences</td>
<td>Session name: Approaches to M&amp;E: MENA experiences</td>
</tr>
</tbody>
</table>

**Institutionalizing results based monitoring and impact evaluation practices in community based social services development projects**

*Presenter: Kristine Grigoryan*

The need for an impact evaluation in the implementation of the development program is important in terms of delivering better results. Impact Evaluation serves as a powerful tool in improving the project design. Moreover, donor agencies should rely on impact evaluation research and evidence in funding competitive grant projects. The best justification for any funded project comes from identified comparison group of participants, which are similar to those benefiting from the project.

In March 2007, the Social Protection System Strengthening project launched its activities in Armenia. One of the crosscutting activities in the project is the grantmaking. The impact evaluation is the key component of grant projects, particularly for the social service delivery projects, which are implemented jointly with state government, local authorities and nongovernmental organizations.

The Community based social services delivery project is based on a wholly participatory and collaborative model, whereby citizens, community councils, local-government bodies, businesses, and NGOs join to identify specific community issues, areas of need, strengths, and weaknesses. With input from all stakeholders, social-service priorities are determined. They may include vocational training, job placement, legal and/or psychological services, and assistance to the disabled, vulnerable senior citizens, and economically vulnerable families. Stakeholders also determine the level and extent of community resources (including funds, volunteer services, facilities, and in-kind services), which the community is capable of contributing to the implementation of social-service delivery. Identified social services are subsequently outsourced to NGOs engaged in social-service delivery, through matching grants. The stakeholders monitor and evaluate the program, and adjust the requirements to
develop sustainable, cost-effective services the community itself will be invested to maintain and institutionalize impact evaluation approach.

The project envisions capacity building for the local stakeholders thus the activities and outcomes are planned jointly with the local government and NGOs. The project includes following three phases:

- First phase of the project is the planning phase, which includes series of community trainings on community SWOT analysis and identification of the social services needs. At this stage the community identifies beneficiaries, how many and what kind of people the project should reach.
- The second phase includes development of the goals, objectives, outputs and outcomes of the project. During this phase each project has well developed logical framework with clear goals, objectives, activities and outcomes.
- The third phase of the project is the implementation phase, where we conduct joint monitoring and evaluation of the project implementation results by the end of the project. Based on the evaluation outcomes changes in the projects, funding, and behavior of the target population.

Through project monitoring and evaluation practices our project is developing the government practice to improve the project. The project builds the capacity of the stakeholders to use the tool of the monitoring and evaluation to strengthen outsourced projects. The whole process of outsourcing leads to development of the logical framework and monitoring of the outsourced projects as well as evaluating outcomes. The project encourages counterparts to use results of the impact evaluations in decision making and funding future projects.

The impact evaluation is a part and parcel of each outsourced project within SPSS. The most important part of the project is local capacity building and emphasis on impact evaluation use for decision making and policy development through outsourcing to civil society players and monitoring and evaluating outsourced projects. The projects support institutionalization of the impact evaluation approach.

| Themes 1: Improving development effectiveness: the role of impact evaluation | Session number: [6] |
| Themes 2: Impact evaluation: Approaches and methods | Session name: Budget support and budget systems |
| Themes 6 : Impact evaluations of policies, programs, projects, partnerships and networks | |

A comparative assessment of the budget transparency: some lessons to be learned for African countries

**Presenter: Zine M. Barka**

“How do we begin to strengthen our ability to assess and improve our government’s budget?” This question is increasingly being asked by members of non-governmental organizations (NGOs) around the world, particularly in African countries or countries new to democracy. The NGOs have realized that their ability to advance their goals — whether these are to combat poverty or to strengthen democratic practices — will be enhanced if they develop a capacity to undertake budget analysis.
This piece of research relies on available experiences and attempts to pull together helpful recommendation to budget work into one coherent document. It explores the systematic overview of the different aspects of effective budget analysis.

To achieve better influence and high policy impact the proposed paper consists of a brief introduction to the stages of the budget cycle and concentrates mainly on how NGOs can contribute at each stage. For conclusion, we offer some fundamental lessons for being effective at applied budget work. These include a discussion of some of the "watchwords" for applied budget work — accuracy, accessibility, timeliness — as well as a discussion of media and dissemination issues and working with policymakers.

Themes 1: Improving development effectiveness: the role of impact evaluation
Themes 2: Impact evaluation: Approaches and methods
Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks

Session stream: Budget support and budget systems

Session number: [6]
Session name: Budget support and budget systems

3 Step approach to evaluation of budget support operations to combine aid and policy, and process and impact evaluation

Presenter: Enzo Caputo
Co-authors: Andrew Lawson, Martin van der Linden

In June 2008, the Joint Evaluation Unit of the European Commission has submitted a new Methodology for Evaluation of Budget Support Operations, to be tested and discussed at the DAC level. The work on such methodology had been assigned to a team composed by Enzo Caputo (Development Researchers’ Network), Andrew Lawson (ODI), and Martin van der Linden (Ecorys) who have produced an Issue Paper, where the 3 Step approach is widely presented and discussed.

Budget Support has become a very important instrument of external assistance to development and it regards, with diversified patterns, different groups of countries, from fragile, to low-income, to middle-income states. Some features identify budget support as a specific aid instrument, including:
- transfer of aid funds to the recipient’s budget
- establishment of a partnership framework between the government and its partners (policy dialogue)
- alignment of the external partners’ support with the governments’ strategies and systems
- harmonisation of the external partners’ support.

Key issues of the 3 Step approach

Any evaluation of budget support needs to assess the actual improvements in the implementation of the government policies supported by the aid instrument. In this respect, it must put a strong emphasis on impact evaluation. On the other hand, especially in the poorer countries, it needs to identify the ways through which Budget Support has supported the government capacity and the contribution given to the results achieved. In this respect it is an evaluation, aiming at improving the aid effectiveness in terms of increased government capacity and mutual accountability between the recipients and their external partners.
Such double priority makes the evaluation of Budget Support particularly tricky. The efforts so far have either focused on the first or on the second assessment. The idea that impact evaluation could not apply to evaluation of Budget Support has been widespread. The 3 Step approach builds on the wide literature and experience on Budget Support evaluation and provides a framework for combining aid and policy evaluation, process and impact evaluation.

Step 1 is typically an evaluation focused on the process of change of the relationship between governments and donors. Its results are very important to improve aid effectiveness on the government systems, and ensure mutual accountability between a government and its development partners, as well as between the latter and their tax payers.

Step 2 is typically an impact evaluation, focused on the actual outcomes and impacts of the sectoral or global government policies that BS has intended to support. Its results are very important to improve the policy effectiveness and the policy focus of Budget Support.

Step 3 allows matching the results of the two previous steps and helps identifying the related links, if any. Its results may help improve aid effectiveness through budget support on

| Themes 1: Improving development effectiveness: the role of impact evaluation |
| Themes 2: Impact evaluation: Approaches and methods |
| Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks |

Session number: [6]  
Session stream: Budget support and budget systems  
Session name: Budget support and budget systems

**Why not more progress towards propoor results-oriented budgeting and monitoring in Sub-Saharan Countries?**

*Presenter: Per A. Eklund*

Poverty in SSA remains high comprising consumption poverty and resilient high non income poverty. The latter reflects limited entitlements and capabilities. Most human MDGs will not be reached for SSA even with high income growth. Shortfalls impact severely the poorest households. ODA resources for SSA can be better used. The 2008 OECD/DAC evaluation of the Paris Declaration of Aid Effectiveness (PD) showed that for 30 SSA countries, achievements to date fall far short of the targets for 2010. Progress has been "alarmingly limited" in developing results-oriented frameworks. Activity- based programming dominates result-oriented budgeting.

The donor originated PD is a blunt instrument for progress. It provides 12 "indirect indicators; they focus on methods rather than on building up capabilities for reaching specific results-based indicators. The PD in formulating indicators did not consider poverty as being multidimensional in concept. Mechanisms do not drive change processes for delivery of cross-sectoral services in line with local priorities.

Stunting in most countries in SSA affects 30% to 40% of the population; it cannot be addressed through compartmentalized policies. Achieving breakthroughs requires secure access to adequate food, sanitary environment, health services and education. Population based monitoring of stunting prevalence at local levels with mechanisms for sharing information with communities may drive demand but is rare (Ashworth et al. 2008). Results oriented cross-sectoral budgeting requires shifts in economic and budgetary planning with support from agriculture, local government, health, water and sanitation, environment, public works and education.
Rethinking the Effect of Cutting Red Tape in Egypt: A Dynamic CGE Analysis

**Presenter: Chahir Zaki**

Improving the efficiency of trade logistics is a critical priority for trade promotion. Hence, relaxing the constraints "behind-the-border" that increase the cost of doing business would contribute much to trade a country’s integration into global trade which, in turn, reinforces economic development. Therefore, Trade Facilitation has become a crucial aspect of trade policy. Actually, the real barrier to trade in many countries, and especially developing ones, is no longer tariffs or quotas, but other impediments such as high corruption, lengthier time to deliver or to clear traded goods, more complicated bureaucracy and poor infrastructure. That is why Trade Facilitation is considered one of the most important Doha’s Round elements that has to be promoted. The "Trade Facilitation" definition adopted here is the one defined in Zaki (2008). This definition splits Trade Facilitation aspects into four major parts: simplification of commercial procedures; harmonization of commercial rules; transparent information and procedures and the recourse to new technologies allowing trade promotion. Three main motivations explain the importance of quantifying such a topic through a CGE model in the Egyptian context: economic interests, empirical reasons and more specifically the Egyptian case.

Regarding economic interests, first, after reducing tariff and non-tariff barriers, trade partners have discovered that there exist other impediments to trade (OECD, 2002a). Reduction of such non-official barriers is likely to have more impact on trade than the reduction of classical ones. Obviously, such impediments involve transaction length, bureaucracy, customs fraud, etc. Second, with the increased supply chains interdependency, imported products delivery delays have turned into a severe constraint on production. This is why customs clearance and delivery of imported products has become a quite important determinant of the production process. Thirdly, the cost of non-facilitation is very high. These barriers account for 2 to 15% of the exchanged goods value. Hence, those economic explanations show to what extent Trade Facilitation process is a crucial issue.

As to the Egyptian case, in 2008, Egypt has been the top reformer in the region and worldwide as it greatly improved its position in the global rankings on the ease of doing business. Yet, Egypt is still ranked 26 overall for Trading Across Borders (Doing Business, 2008) but it is much better that many other comparator economies such as Turkey, Lebanon, Jordan or Syria. In addition, The World Economic Forum issued its first “Global Enabling Trade Report” in which Egypt ranks a low 87th for the ease of getting goods across the border. The report underlined the positive as well as the negative aspects of Egypt’s Trade Facilitation aspects.
### Modeling Gender Dimensions of the Impact of Trade Liberalization in Pakistan

**Presenter: Rizwana Siddiqui**

The paper uses a computable general equilibrium model specially constructed for investigating gender dimensions of the effects of trade liberalization in Pakistan from both production and consumption sides of the economy. The model employs various indicators to measure the impact - monetary poverty (Foster, Greer, Thorbecke, or FGT, Indices), time poverty (leisure), capability poverty (literacy and infant mortality), and welfare effects (Equivalent Variation or EV). The simulation results show that revenue neutral trade liberalization increases employment of unskilled labor of women and increases women’s real wage income more than men for all types of labor but keeps division of labor biased against women. It adversely affects women in relatively poor households by increasing their workload, deteriorating capabilities and increasing relative income poverty. However, these effects remain gender neutral or favor women in the richest group of households-high education and employer.

### An impact analysis of a home-visiting ECD program in the Caribbean

**Presenter: Wendy Janssens**

Child development during the early years of life lays the foundation for future life outcomes. Especially children from poor families are at risk of falling behind their more fortunate counterparts. Early developmental gaps are difficult to close at a later stage. Whereas a large body of literature documents the potential beneficial impact of Early Child Development (ECD) interventions in industrialized countries, quantitative evidence for developing countries is still relatively scarce. The existing impact studies focus mostly on the evaluation of center-based instead of home-based interventions.

This research evaluates the one-year impact of a home-visiting ECD program in the Caribbean that targets vulnerable children aged birth to three and their caregivers. The aim of the program is to change inappropriate parenting practices in order to enhance healthy child development. The intervention consists of weekly facilitator visits to the child and caregiver in the home as well as monthly parent meetings at local community centers.

The research is set up as a longitudinal, quasi-experimental study that follows approximately four hundred children over a period of three years. The study is conducted in fifteen low-income communities in the Eastern Caribbean island St Lucia. Half of the children live in
communities where the program was introduced after the baseline survey in 2006. The other half live in matched control communities. An analysis of child developmental status at baseline shows that the treatment and the control group were highly comparable before the introduction of the intervention.

In each survey round, the children participate in a psychological assessment of their cognitive and socio-emotional development. Their caregivers are interviewed using an extensive parenting questionnaire. The econometric analysis is based on a panel of two survey rounds. It uses differences-in-differences and propensity score matching methods to estimate program impact based on an intention-to-treat approach.

The results are unexpected. One year after the introduction of the program, we do not find a significant impact of RCP on any of the cognitive or socio-emotional summary and sub-scales. Examination of treatment heterogeneity suggests that the program has a small positive impact on fine motor skills and socio-emotional development of those children who enter the program at a younger age, as well as on the gross motor skills of children who were enrolled for a longer period of time. We find no further evidence of differential treatment effects.

The most dramatic finding is that children who enroll in daycare centers unrelated to the intervention program score significantly higher on several cognitive development subscales. In both treatment and control communities, enrollment in center-based daycare has taken up substantially over the past two years. In fact, half of the children who prematurely dropped out of the intervention are currently enrolled in daycare. We discuss implications for our evaluation design and results.

**Theme 2: Impact evaluation: Approaches and methods**

**Theme 6: Impact evaluations of policies, programs, projects, partnerships and networks**

**Session number: [8.1 ]**

**Session stream: Child health and development**

**Session name: Child health and development**

**Experiences with Lot Quality Assurance Sampling in Save the Children health and food security programming in Uganda**

**Presenter: Luc F.E. Vanhoorickx**

As part of its broad program portfolio, Save the Children in Uganda (SCiUG) is implementing a 5-year Food Security program and a 4-year Community-Based Family Planning Project in the central region of Uganda. For these programs, SCiUG carries out quantitative household surveys using Lot Quality Assurance Sampling (LQAS) to monitor the program’s higher level results at the population level. The goal of the surveys is to assess results at the impact level, as well as towards outcomes such as the adoption of promoted health/nutrition practices or agricultural techniques. Indicator values are comparable to baseline values, previous years’ data and final targets.

This presentation will first briefly present both SCIUG programs and the LQAS methodology, including its advantages over other methods such as cluster sampling and share the Lessons Learned. In SCIUG’s programs, LQAS surveys employ a two-step random sampling procedure where communities for the survey are selected in a first step using sampling proportionate to size of the community, and then the households in the communities are selected in the field using random tables to pick from a list of households, “spin the bottle” technique and/or random GPS coordinates. Data collection tools are adapted from standard survey questionnaires for the outcome and impact indicators of the programs. In some areas, these
were transferred to Personal Digital Assistants (PDAs) to pilot the use of these devices in the survey work. SCIUG extension workers carry out the field work; while this calls for mixing them up to eliminate interviewer bias, it also keeps the surveying capacity in the organization to allow for annual LQAS monitoring exercises.

Some of the advantages of using LQAS are that the methodology not only permit smaller sample sizes than traditional random sampling methods, but also allows disaggregated analysis at the local level, for management decision making. Indeed, the impact area of the program in the central region of Uganda are divided in so called supervision areas or ‘lots’ and with LQAS, the conclusions for all indicators can be disaggregated by supervision area.

Some Lessons Learned include that due to the smaller sample size, not much deeper statistical analysis can be done with the obtained data (in contrast to for example cluster samples). However, indicators are disaggregated by supervision area and further analysis such as splitting up parts of the indicators can be done. Then it should also be noted that even though sample sizes for the LQAS methodology are smaller than for other sampling methodologies, this does not necessarily translate into a drastic reduction of the number of person-days needed for the field work because the sampling points are randomly distributed throughout the supervision areas. Also worth noting is that for deeper evaluation of the impact of the programs, qualitative research is used to explain the results found with the LQAS surveys. Indeed, LQAS is ideal for identifying what the problem is and where in the program impact area it is most pronounced, leading the qualitative research in these areas to investigate why that is the case.

Themes 2: Impact evaluation: Approaches and methods

Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks

Session number: [8.2 ]

Session stream: Child health and development

Session name: Child health and development

Madrasa Early Childhood Development Community Based Monitoring and Evaluation (CBM&E) System

Presenter: Wayne A. Bacale

Madrasa Resource Centre Early Childhood Development programme in East Africa is a shining example of a programme producing young children “shinning stars” at preschool level and beyond. The Madrasa ECD Community based Monitoring and Evaluation is a system that uses participatory methods and tools to provide relevant information for decision making at the community and programme level, generate lesson learned and best practices. The primary stakeholder at the preschool is the Community Resource Team, with the support of Teachers and management intervention from the School Management Committee. At the cluster level, the prime movers are the Cluster CRT leaders with the support of the preschool CRT members.

An ongoing mentoring and training to operationalize CBM&E is crucial along the process while clearly defining the M&E system. As a start, we have built the capacity of fifteen Training of Trainers (composed of Community Leaders) that rolled out the CBM&E training at the community level. Furthermore, ongoing mentoring and on-the-job training on data collection, analysis and feedbacking is being done by the TOT and MRC staff. This system has provided substantial information for decision making and at the policy level. As an example, the child tracking system provided evidences on the learning development of children. Furthermore, most of the tools provided above provided the opportunity of transparency and accuracy of data.
Children affected by AIDS: Strengthening families through income support and family-focused services

*Presenters: Lorraine Sherr, Michelle Adato, Chris Desmond*

The Joint Learning Initiative on Children and AIDS (JLICA) was launched in 2006 with the aim of reviewing and synthesizing available evidence on the nature of the problems to be addressed and promising interventions for effective responses at the national level in countries hard hit by the epidemic. Over two years, more than 50 research reports were produced and thousands of stakeholders were involved in discussions through conference presentations, website distribution, newsletters, and other forms of dissemination. The work of the JLICA Work was conducted through four learning groups – Learning Group (LG) 1 on Strengthening Families; LG 2 on Community Action; LG 3 on Expanded Access to Services, and LG 4 on Social and Economic Policies. Learning Group 1 framed the following key JLICA findings:

1. HIV transmission in high prevalence countries is concentrated within families and occurs mainly through partner and parental transmission. Similarly, the impact of AIDS is felt most intensively at the family level, and the family is the most important social network for both prevention efforts and mitigation for children.
2. Interventions must target all poor children in heavily HIV affected communities, not only children infected with HIV and those orphaned as a result of AIDS, because all already vulnerable children feel the shockwaves of the epidemic on their communities and the services they receive.
3. Children’s wellbeing can be improved through policies and programmes directed at strengthening families, as well as essential services such as health care and education, because families provide care and protection for children on a daily basis as well as over their lifetime.
4. National social protection policies are a critical lever in family strengthening to improve the care and protection of children. Social protection includes employment and livelihoods, the social wage, and social security.
5. Direct income transfers are the most effective and efficient means of improving the health, development and education of vulnerable children. Ensuring that money reaches vulnerable households directly has immediate and long-term benefits for children.
6. HIV/AIDS prevention, treatment and care services should attempt to provide integrated care to the whole family. In this paper, we outline the methods of systematic review and integration used in the JLICA, and the evidence on which the key recommendations are based. The presentation concludes with an overview of the impact the JLICA findings are having on policy, programmes and practice in this field.
Impact Evaluation of Child Survival Program in Sierra Leone

Presenter: Yuki Suehiro

Co-authors: Bockarie Sesay, Sowo Lebbie, Sayoh A. Francis

CARE Sierra Leone and Ministry of Health and Sanitation (MOHS) implemented a five-year (2003-2008) USAID-funded Child Survival Project to improve the health status of children under five and women of reproductive age in five chiefdoms of Koinadugu District in Sierra Leone. Project objectives were: 1) Strengthened family and household knowledge and decision-making skills related to health of women and children resulting in the practice of positive behaviors to prevent, recognize and manage common diseases; 2) Enhanced community capacity to form groups and institutions that sustain health initiatives, demonstrate social cohesion, and promote good governance mechanisms; and, 3) Improved quality and accessibility of services provided by MOHS personnel and MOHS extension services.

Intervention focuses included: Expanded Programme on Immunisation; Nutrition; Malaria; and Maternal and newborn care. The project organized the community into community health clubs (CHC); trained them in maternal/child survival and general health issues; promoted community led action and participation in the delivery of health services. Furthermore the project worked with Village Development Committees to build their capacity to monitor and support community utilization of health services. The project also worked on the health system strengthening through capacity building of service providers and quality assurance approach.

In August 2008, the project final evaluation was conducted to measure the impact of the intervention and to draw lessons learned for future child survival programming. Measuring sustainability was one of the key focuses, for which child survival sustainability assessment (CSSA) methodology was applied. The evaluation involved three-step process:

- Quantitative: A standard KPC 2000+ 30x15 cluster survey (including anthropometric measurement) was conducted.
- Health facility assessment: COPE (Client Oriented Provider Efficient) Assessment was done in collaboration with MOHS.
- Qualitative: Participatory qualitative assessment was done which involved an external consultant, stakeholders and partners.

Key findings include:

Nutritional status – percent of underweight children aged 0-23 months came down from 26.5% (baseline) to 18.8% (endline).

Knowledge and behaviors – significant improvement in knowledge and behaviors related to child and maternal health was observed such as: early initiation of breast feeding (baseline 20% vs. endline 70%); exclusive breastfeeding among 0-6 month infants (8% vs. 68%); children 0-23 months sleeping under insecticide treated bed nets (1% vs. 83%); women of reproductive age who knew at least two obstetrics danger signs needing referral (38% vs. 75%).
Access to services and quality of care - the project achieved significant improvements in quality of services; availability of staff; improved logistics and drug availability; network for emergency referral; strengthened accountability; and most importantly community participation in formal health system.

Local capacities – immense amount of community cohesion and ownership was observed, which ensures sustainability of the positive health behaviors. Community participation has matured beyond health and addressed broader developmental issues as well.

This paper discusses further in depth: 1) findings of the evaluation with analysis of the determinants of success, constraints and other factors affecting project performance; 2) best practices and overarching lessons learned from the project; and 3) contribution of participatory impact evaluation methodology to an effective learning process.

Using the Most Significant Change Technique for actual program improvements

Presenter: Awni Amer

Introduction. Plan is an International non-governmental development organization has been in existence since 1937. Plan’s vision is a world in which all children realize their full potential in societies that respect people’s rights and dignity. The Impact evaluation is one of the core components of the new approach that Plan adopts. Plan started to use its new Monitoring & Evaluation (M&E) framework called Program Accountability and Learning System (PALS). To improve the impact evaluation, the new system strengthening Plan’s accountability to the children, communities and partners, continually improve program quality and effectiveness through learning process and finally supports continuous learning about Rights-Based Approach that can be shared within and outside Plan.

Background. Plan is shifting from a needs-based to a Rights-Based Approach that adopts participation in all the project/program phases, Plan started to use its new M&E approach called Program Accountability and Learning System (PALS). The new M&E system is giving more attention to involve all the stakeholders and partners in the evaluation process to get their feedback through community reflection processes. One of the new participatory tools used widely now in Plan Egypt is the Most Significant Change (MSC) technique to promote downward accountability, sharing and collective community learning.

MSC Technique. The MSC technique is a form of participatory M&E tools developed originally by Dr. Rick Davis & Jess Drat 2005. It involves many stakeholders in deciding the sort of changes to be considered. It occurs throughout the program cycle and provides information data on impact and captures outcomes that can be used to help assess the performance of the program as a whole and do the needed improvements. This technique is managed through 10 steps to collect/select the stories from the stakeholders/beneficiaries during /after the project life, selecting them through different levels (sub-community, community, Program unit, and national level).
Plan Egypt is using MSCT widely and trying to use the feedback resulting from the application of MSCT facilitated and managed by the community volunteers (M&E teams). The following are some of the recommendations resulting from using the MSCT to capture the outcomes of two of the Female Genital Mutilation (FGM) projects:

- Mobilizing men against FGM is a very important intervention for the vital role that they may play in discouraging FGM.
- It is recommended to include the basic life skills intervention in the awareness component provided for all ages starting at 9 years.
- Using the negative models as result of the practice of the FGM played a vital role in saving others who are still at risk of the FGM.
- Support the role of the school, mosque, children clubs and church should be continued and enhanced and promoted on different levels in the project design.

Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks

Session number: [9]

Session stream: Decentralisation

Session name: Decentralisation

Does transparency make local governments more responsive? Evidence from the Philippines using difference-in-difference approach

Presenters: Maria Melody S. Garcia Joseph J. Capuno

Co-author: Joseph J. Capuno

Increasingly, many developing countries like the Philippines are adopting fiscal decentralization policies to improve the frontline delivery of public services. The underlying premise is that local officials are directly accountable and therefore more responsive to service clients. Local government officials can be made accountable if, among other things, the people have access to the appropriate and timely information. This paper tests the proposition that transparency make local governments more responsive, using data from a local governance project piloted in four Philippine provinces in 2001-2003.

In this local governance project, an index of local government performance was introduced in twelve cities and municipalities. Eight of the twelve sites are designated as treatment areas where the scores in the index were announced, through posters, leaflets, komiks (popular magazines) and public presentations, in two successive years. In half of the treatment sites, the local government themselves generated and disseminated the scores, while in the other half, civil society organizations carried out the same activities. The rest of the sites are designated as control areas where the scores were not announced. To measure the impact of the scores on the responsiveness of the local government officials, a baseline household survey was conducted, followed by another two household surveys after each round of the score announcements. The dataset of 3,600 household respondents comprises information on the responsiveness local governments and other outcome variables in both treatment and control areas, and before and after the index was introduced.

To evaluate the impact of the index on the responsiveness of local governments, a difference-in-difference approach is adopted here. Specifically, the differences in the responsiveness of the local government officials before and after the introduction of the index in the control sites is estimated, and then subtracted from the differences in the responsiveness of the local government officials before and after the introduction of the index in the treatment sites. Thus, the double differencing is able to eliminate both time- and site-invariant factors that
condition responsiveness. Since there are different respondents in each survey round, individual-level covariates are also introduced in the regressions to control for other factors.

The difference-in-difference estimates show that the score has positive and statistically significant impact on the overall responsiveness of local government officials, but a negative and statistically significant impact on the responsiveness of the local chief executive (mayor) to complaints. Further, the impact varies with the type of local agent (i.e., local government or civil society organization) that makes the announcement. Overall, the results thus show that public disclosure of local government performance can influence the behavior of local government officials, who for better or worse are now assessed more critically by their constituents. Hence, measure to improve transparency can help lead improvements in local public services under decentralization.

<table>
<thead>
<tr>
<th>Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks</th>
<th>Session number: [9]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session stream: Decentralisation</td>
<td>Session name: Decentralisation</td>
</tr>
</tbody>
</table>

**Evaluating the Impact of Decentralizing Courts in Peru**

**Presenter: Raul F. Andrade**

This study seeks to provide empirical evidence for the theoretical presumption about the role that an efficient judiciary has for economic growth. At the micro level, having an efficient judiciary is needed to enforce contracts, facilitate economic transactions and regulate social interactions. The specific objectives of this study are to evaluate if increasing access to justice has positive effects on: i) the frequency of conflicts, such as: crimes, unpaid debts, unpaid child alimony, etc., ii) the rate of access to formal versus informal means to solve controversies, and iii) the efficiency of the conflict resolution mechanisms in terms of time and cost. Overall, by providing information on each of these aspects, it will assess if the presence of formal justice may have a positive impact on wellbeing.

For these purposes the impact evaluation of a program to decentralize courts in Peru (Basic Modules for Judicial Services Provision -BMJ program) will be pursued. Through the BMJ program 43 new courts in localities of Peru with a lack of judicial services were implemented. Primary data on a sample of 2,500 household living in localities with BMJ and localities without BMJ has been collected. The survey questionnaire asks for information about the incidence of the abovementioned conflicts (frequency of conflicts), the mechanisms that people use to solve these controversies (access to formal and informal justice), and the costs in time and money that solving these controversies implies (transaction costs).

Two identification strategies are used to obtain causal evidence on the effects of the BMJ program. First, propensity score matching was used to create a sampling framework. The probability of a locality being a beneficiary of the BMJ program was estimated using pre-treatment characteristics. Geographic, demographic, and economic characteristics of the localities, as well as variables related to the availability of social services were included to estimate the propensity to receive the treatment. The propensity score was then used to construct a list of twin localities: a list of matches between a treated locality and nearest twin from the set of control localities. The list of twin localities was the sampling framework from which the random sample of 2500 households was chosen. Among this sample, the group of households living in localities with a BMJ and the group of households living in localities without a BMJ are similar on average. The impact of living on a locality with a BMJ on the outcomes abovementioned will be found by means of regression analysis.
The second identification strategy is based on the presumption that the final places where BMJ were constructed were chosen on the basis of political considerations. President Fujimori had the tendency to favor localities that he considered important for his electoral ambitions by providing them with different services, such as judicial services. This leads to the opportunity for using an instrumental variable method, modeling the construction of a court as a function of electoral variables at the locality level, and then using the variability that this instrument generates on the sample of households to identify the differences in the abovementioned outcomes.

### Themes 1: Improving development effectiveness: the role of impact evaluation

<table>
<thead>
<tr>
<th>Session stream: Development agency support for impact evaluation</th>
<th>Session name: Development agency support for impact evaluation</th>
</tr>
</thead>
</table>

#### A comparative study of Evaluation policies and practices in development agencies

**Presenters: Martin Prowse, Foresti, C. Archer, T. O’Neill and R. Longhurst**

This report aims to understand experiences and challenges in implementing development evaluation reform and to assess a variety of solutions, looking at different development institutions. It hopes to lead to improved knowledge of evaluation systems in different development agencies and to provide some pointers to support AFD and other development agencies in their evaluation reform processes.

An overview of the evaluation units in the institutions under study looks at: accountability; information and knowledge; partnerships; capitalization of experience; and feedback. The report compares policies and practices across institutions, looking at different choices and their consequences, as well as strengths and weaknesses. The analysis also identifies the main internal and/or external factors involved for each institution. The report reviews management arrangements and structural position of evaluation units; their main features, processes and tools; and practices involved in commissioning, managing and supporting evaluation processes. It does not aim to identify good practice or to set standards for the AFD to follow. The report is based on full case studies (including primary data collection) of five evaluation units, at: DFID; SIDA; the World Bank; KfW; and the AfDB. Staff for interviews mainly came from the evaluation unit. It also covers four desk-based studies (relying mainly on documentary analysis), at: the EU; the IMF; Danida; and Oxfam. In addition, a small number of key informant interviews were carried out with leading development evaluation experts.

### Themes 1: Improving development effectiveness: the role of impact evaluation

<table>
<thead>
<tr>
<th>Session stream: Development agency support for impact evaluation</th>
<th>Session name: Development agency support for impact evaluation</th>
</tr>
</thead>
</table>

#### Post intervention Studies

**Presenter: Junaid Habib**

**Aims and objectives.** The overall aim of post-intervention studies is to document the lessons learned and to collect evidence towards Plan’s contribution to the long-term changes and results brought about through Plan’s interventions. Each post-intervention study will have its context specific objectives including analysing and documenting the lasting results (positive / negative), lessons learned and factors of sustainability.
**Methodology.** For the proposed studies, we are exploring the methods which show rigour and comparative strengths. Recognising the fact that the long term changes cannot be attributed to a single actor / organisation, rather it can only be realised through the contribution of a number of actors and factors, the methodology to conduct these studies will be based on mixed methods *(which are defined as the explicit adoption of both quantitative and qualitative methods in the evaluation design).*

Based on the background desk review quasi-experimental designs offer more rigour and design strengths over traditional impact assessment methods although no single methodology will be imposed on the evaluators to allow creativity and innovations. For the proposed post-intervention studies a credible counterfactual analysis could be developed using a comparison / control group. The comparison group could be a neighbouring community or a physically distant community but with identical socio-economic indicators and comparable characteristics. While determining the comparison groups we should be aware of the ethical concerns involved, if any.

The two important design elements for the estimation of project impacts are: a carefully selected control (or comparison) group and pre-and-post intervention comparison of the two groups. The first element is essential for the formulation of logically sound counterfactual *(the comparison of pre-test and post-test control group is the counterfactual – what would have happened to the project population if the project had not taken place).* While the second element, observation at two points in time, is usually desirable, it is possible to use post-intervention project and control group cross-section data only.

<table>
<thead>
<tr>
<th>Themes 1: Improving development effectiveness: the role of impact evaluation</th>
<th>Session number: [10]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session stream: Development agency support for impact evaluation</td>
<td>Session name: Development agency support for impact evaluation</td>
</tr>
</tbody>
</table>

**Impact evaluation of project interventions targeting child labour: a case study**

*Presenter: Peter E. Wichmand*

An impact evaluation initiative involving ILO/IPEC and the Understanding Children’s Work (UCW) would be presented as a case study of impact evaluation in the area of child labour. The impact evaluation initiative, begun in January 2009 and lasting for 52 months, is aimed at extending the evidence base on the impact of program interventions targeting child labour. Through a combination of methodological development, targeted field research, capacity building and research dissemination, the initiative will contribute to building and spreading knowledge on the impact of child labour interventions and ways to measure such impact. Robust empirical evidence relating to the impact of policies and programs on child labour remains limited, and the initiative will help to fill this information gap.

The core component of the initiative involves implementation of impact evaluations of three specific program interventions directly addressing child labour. In the first of an envisaged two-stage research effort, this project component will be directed towards not only generating initial data, but also, and even more importantly, at establishing replicable "protocols" and standardised instruments for impact evaluation of various types of child labour programs, including evaluability frameworks. Impact evaluation will be included in the projects from the design phase and this initiative will therefore be able to have a degree of influence on the type of interventions to be designed to facilitate credible impact evaluation. The second stage of the evaluation process, beyond the scope of the current initiative, would focus on follow-up data.
collection for the programs included in this evaluation exercise, allowing for a comprehensive evaluation of both short- and longer-term program intervention impact.

A second component of the initiative involves support to integrating child labour modules into planned impact evaluations by other organizations of programs in relevant social policy areas. Child labour is a complex phenomenon that cuts across policy boundaries – education, health, labour markets, basic services, capital markets and social security, inter alia, all play an important role. Yet quantitative impact evaluation instruments applied to programs in these policy areas rarely include child labour as an explicit outcome variable, or, if included, rarely accord it relevance, meaning that little is known about the contribution of the programs to child labour reduction. Technical support will therefore be provided through this component to adding questions or modules on child labour to planned impact evaluations of programs with a bearing on child labour.

The proposed presentation to the Cairo conference will discuss these and other elements of the impact evaluation initiative, as a concrete case study of impact evaluation in the area of child labour. The presentation will cover some of the technical elements of the initiative, as well as how the initiative fits within broader impact assessment efforts undertaken by ILO-IPEC and other implementing agencies. It will also look at how findings from the initiative will be “mainstreamed” into on-going program development and impact evaluation activities.

**Themes 2: Impact evaluation: Approaches and methods**

**Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks**

**Session stream: Education**

**Session name: Education**

**Free Primary Education Programme in Kenya: An Impact Evaluation Using Propensity Score Methods**

**Presenter: Milu Muyanga**

**Co-authors: John Olwande, Esther Mueni**

Lack of formal education is a medium through which poverty is passed on from one generation to the next. Missed schooling or poor educational achievement by children from poor parental background can undermine their chance to escape from poverty in the future. Government intervention can lead to enhanced access to education by children from poor households. For sometime, primary education system in Kenya has been characterised by high wastage in form of low enrolment, high dropout, low completion and poor transition rates and grade repetition. This scenario was attributed to high cost of primary education. To reverse these poor trends in educational achievements, the government initiated free primary education programme in January 2003.

This paper evaluates the preliminary impact of FPE programme in Kenya to assess whether the programme is succeeding in reversing poor education trends using panel data. First, correlates of selected primary schooling outcomes are established using regressions with a step dummy to capture the impact of the FPE programme. The impact of FPE programme is again evaluated using propensity matching methods. Lastly, benefit incidence analysis is conducted to establish the progressivism of the FPE programme.

The regression results show that while primary and secondary enrolment improved with the introduction of FPE programme, primary school grade progression declined. The same results
also emerge from the propensity score matching analysis. The primary school enrolment momentum set in motion after the introduction of FPE education is being sustained.

The results from the benefit incidence analysis show that government spending on FPE programme is progressive. The 20 percent poorest households capture more than twice the benefits of FPE captured by their counterparts in the 20 percent wealthiest income group. Even though the programme was not targeted, by default poorer households happen to have more children. Consequently, the poor are likely to have more children enrolled in primary schools when financial setbacks are alleviated. Thus, this result is more or less a function of household behaviour rather than government behaviour. However, the poorer households have a potential to benefit a lot more from the programme if the children out of school could be enrolled. More enrolment of school going aged children from poor families that are still out of school would require pragmatic interventions to combat other factors beyond direct schooling costs that keep them from enrolling for schooling. Such interventions would definitely require an inquiry into the relevant hindrances to primary school enrolment, before they can be instituted.

Assessing the Impact of Free Primary Education Policy on Access and Schooling Outcomes in Kenya

**Presenter: Moses Ngware**

Kenya introduced a free primary education (FPE) policy in 2003 that was implemented in all public primary schools. This provided an opportunity to assess factors, other than lack of fees, which affect enrollment and retention in schools. The policy enabled children who were out of school primarily due to lack of school fees and other levies to return. For those who enrolled, the challenge is to ensure completion of primary education and transition to higher levels of schooling. To assess the impact of FPE on schooling outcomes, the African Population and Health Research Center (APHRC) initiated a longitudinal study among slum and non-slum residents of Nairobi. In addition, the study explores timing and linkages between school enrollment, retention and dropout, and onset of risky behaviors such as early initiation of sexual activity, unintended pregnancy, smoking, alcohol and drug abuse, and gang membership.

The slum population is derived from two informal settlements where the APHRC is currently implementing a longitudinal health and demographic surveillance system. A middle income non-slum area is covered in order to establish the relative impact of slum residence on the linkages between school enrollment, retention and dropout, and the onset of risky behaviors in the wake of FPE. The study tracks children aged 5-19 years old. For the older children, the study provides an indication of factors associated with transition to secondary school in the context of FPE.

For all those who dropout, in-depth interviews are conducted to assess reasons and circumstances of the dropout. In addition, information on school-related factors are collected from the schools that serve children in the study communities. The key outcome of the study is to assess the impact of FPE on schooling outcomes among the urban poor. This will help identify feasible interventions to improve school participation and transition to secondary
school in poor and middle income urban communities. The three research questions that guide this study are:

1. What is the impact of free primary education on school enrollment patterns and dropout rates among urban slum and non-slum children?
2. What factors are associated with school participation (enrollment, attendance, repetition, performance and dropout) among urban slum and non-slum children?
3. What are the (causal) linkages between school participation and the onset and extent of indulgence in risky behaviors in children?

While the policy appears to have had a positive impact on school enrollment, its initial assessment indicates that some children did not enroll in schools and some who enrolled initially have since dropped out. Furthermore, the benefits of the new policy are inequitably distributed across social groups. Overall, Nairobi City’s poor performance in enrolment has been attributed to the large proportion (60%) of the city’s residents who live in slum settlements characterized by poor schooling facilities, low school attendance, and high prevalence of risky behaviors.

Themes 2: Impact evaluation: Approaches and methods
Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks
Session stream: Education
Session number: [11.1]
Session name: Education

Evaluating the Impact of the Capitation Grant Programme on Basic Education in Ghana

Presenter: Dominic Pealore

The implementation of the various education policies in Ghana since the turn of 2000 was enshrined in the Education Strategic Plan 2003-2015.

Nevertheless its development was guided by the 1992 Constitution provision of Free Compulsory Universal Basic Education, the Ghana Poverty Reduction Strategy, Education for All, the Millennium Development Goals, New Partnerships for Africa’s Development as well as Education Sector Reviews

One of the ten Policy Goals of the Strategic Plan is Increase Access to and Participation in Education and Training.

Efforts were made by way of enrollment drives and sensitization of communities to send their children to school but the fcube had not made the expected impact by 2004 two years before the expiry of its implementation (Primary GER was 86.33% nationally in 2003/04 and 76.06% for deprived districts)

The Capitation Grant Scheme. Despite the instruction that non-payment of fees should not injure a child’s chance of attending school, yet the charging of fees negatively affected attendance at basic schools.

Consequently Government abolished all forms of fees/levies and in place introduced the per capita allocation to every basic school.
Monitoring for impact in a programme’s sphere of influence – a case study of the quality education and vulnerability programme in Zimbabwe

*Presenters: Madri Jansen van Rensburg, Amorlett Nyamweda*

HIV prevention projects use peer education as a participatory method. This is used across age groups and in different settings (including using the workplace as entry point to communities). In many instances unexpected outcomes are reported including developmental changes in communities beyond HIV prevention. This review investigates the planned and unexpected changes that occur when using peer education. The review includes studies from different Southern African low income countries (including South Africa, Malawi and Zambia, amongst others). Studies reviewed include both quantitative and qualitative and mixed method approaches.

Assessing the Impact of Ishraq, a Second-Chance Program for Out-of-school Rural Adolescent Girls in Egypt

*Presenters: Ghada Barsoum*

*Ishraq* is a second-chance program that targets out-of school adolescent girls in rural Upper Egypt. Taking a holistic approach, Ishraq’s curriculum emphasizes literacy and life skills with special attention to reproductive health, livelihood skills, civic engagement, and sports. Girls who pass the literacy exam are allowed the opportunity to be mainstreamed to the national education system. The program was first piloted by four international non-governmental organizations (the Population Council, Save the Children, the Center for Development and Population Activities (CEDPA) and Caritas) in partnership with local intermediary non-governmental organization. The pilot phase reached 277 girls ages 13-15 in four rural villages in the Upper Egyptian governorate of Minia. Building on the evidence-based and documented positive impact of the pilot experience, *Ishraq* was replicated twice from 2004 to 2007. In its current phase, *ISHRAQ* is transitioning from an innovative pilot to a larger program that is currently being institutionalized through the National Council for Youth. In this current transitional stage, *Ishraq* will be implemented in 50 rural Upper Egyptian villages, aiming to reach 2,800 out-of-school girls aged 12-15. A major objective of this phase is to execute an effective scale-up strategy where technical and project expertise will be transferred from the international partners to the implementing and coordinating body of the National Council of Youth at the central, governorate and village levels.

The benefits of monitoring, evaluation and impact assessment have been recognized by the partners implementing the *Ishraq* program from the start. A monitoring and evaluation
component, which made use of both qualitative and quantitative methods, was incorporated into the pilot phase of the program. The results of these evaluation efforts are reported in Brady et al. (2007). In the current expansion phase, impact assessment uses a quasi-experimental design, including baseline and end-line surveys in randomly selected control and intervention villages. The surveys are administered to a panel of out-of-school adolescent girls in both types of villages. The impact of the program is assessed by comparing end-of-program responses to baseline responses. The specific areas in which we assess the impact of the projects are: (i) functional literacy, cognitive skills and continuing schooling, (ii) attitudes about marriage and childbearing, (iii) knowledge about nutrition, hygiene, and reproductive health, (iv) attitudes about harmful traditional practices, such as female genital mutilation/cutting, (v) social isolation, peer networks, and participation in group or community activities, and (vi) gender norms, gender-based violence, mobility, and agency, including attitudes about participation in sports and recreation activities.

This paper will discuss the lessons learned from the impact evaluation process during the pilot phase and the methodology being adopted for impact evaluation in the current expansion phase. Specifically, the paper will focus on ways to combine qualitative data, quantitative impact evaluation data from baseline and end-line surveys, and process monitoring data such as participation and attendance levels. The paper will also focus on pre-analysis activities for research in the expansion phase such as intervention/control village selection procedures; questionnaire design issues; processes for program recruitment and community mobilization activities.

Themes 2: Impact evaluation: Approaches and methods
Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks

Session number: [11.2]
Session stream: Education
Session name: Education

Programme De L’education De La Petite Enfance À Madagascar

Presenters: Jocelyn Ratsiferana

Nombreux sont les rapports des organisations internationales qui insistent sur l’éducation dès le plus jeune âge. Ils considèrent que c'est pendant la petite enfance que doit avoir lieu la première phase du processus de formation tout au long de la vie. Selon eux, cette période de l’existence doit occuper une place essentielle dans un programme d’action éducative, sociale et familiale efficace.


Il est décidé qu’avant de passer à l’échelle, d’expérimenter les activités de l’éducation parentale et du préscolaire au niveau national. Ainsi, il est proposé de faire un état des lieux pour plusieurs raisons, d’une part, pour obtenir une photographie de la situation de base pour pouvoir évaluer les effets positifs de l’intervention par la suite, et d’autre part, pour appuyer la validation des profils (les repères du développement) de l’enfant malgache, développés depuis 2005.
Recueillir et analyser la situation en ce qui concerne le développement psychomoteur, cognitif, langagier, socio-affectif des enfants de 3 ans et de 5 ans 11 mois, ainsi que les pratiques en matière de soin et encadrement.

Cette analyse de situation de base sera utilisée à trois fins : 1) déterminer l’orientation des interventions d’éducation parentale et du préscolaire en mettant à la lumière les pratiques positives et les pratiques négatives en matière de soin et d’encadrement des jeunes enfants; 2) établir une « Baseline » en vue de mesurer l’impact des interventions en matière d’éducation parentale et le préscolaire communautaire dans le futur ; et enfin, 3) finaliser les profils (repères) de développement de l’enfant malgache à 3 ans et 5 ans 11 mois.

Comme l’échantillon que l’on projette de constituer doit être représentatif de l’ensemble de la communauté rurale malgache et pour s’assurer de la représentativité de l’échantillon au niveau régional ou national, on a constitué des zones géographiques représentatives recouvrant assez bien le pays comme le montre la carte ci-dessous.


Themes 2: Impact evaluation: Approaches and methods
Themes 6: Impact evaluation of policies, programs, projects, partnerships and networks
Session number: [11.3]
Session stream: Education
Session name: Education

Good education policy: Lessons from randomized evaluations

Presenters: Rebecca Thornton, Kamilla Gumede

Education is widely seen to be critical to development. Progress is being made towards universal primary school enrollment and we now know a lot about how best to improve access to education. The quality of education children receive through much of the developing world remains poor. Teacher absenteeism rates are high, curricular are often inappropriate, and studies show that many children who attend school regularly are unable to read simple sentences by grade 3.

Over the past 10 years, J-PAL researchers have or are conducting 29 randomized evaluations on education, 12 of them in Africa where progress towards the Millennium Development Goals has lagged behind global trends. Our research cuts across issues of access, quality, and governance. The result is a substantial evidence base on what constitutes good education policy. It is of direct relevance to practitioners involved in developing educational strategies and researchers who want to know how randomized evaluations can be used to study education programs.

The presentation will provide an overview of the common lessons that arise from these multiple randomized evaluations and discuss the extent to which results in one country setting replicated to other settings. It will include findings on: i) the most effective ways to increase school participation; ii) class size effects; iii) teacher accountability; iv) student, parent and teacher incentives; to v) the benefits of getting teachers to adapt teaching to the level of the students. The focus will be on primary education as this is where there is most evidence and considerable policy interest but we also touch on ongoing work to assess the returns to secondary schooling in Ghana.
Effective Rehabilitation and the Role of Evaluation

Presenter: Sajjad Akram

Pakistan has a sturdy Public Sector Development Program (PSDP). There were 2119 PSDPs amounting Rs 335 million, spread over the country, during 2007-08 financial year. The programs are approved, monitored and evaluated by the Planning Commission (PC) of Pakistan. The PC is directly under the supervision of Prime Minister of Pakistan and has nine technical members. Level problems and constraints of Development of Country are often equated with the efficiency and effectiveness of its public sector organizations. Thus, building public sector capacity and performance has become key objective in development. Accessing public sector organization performance is not as easy as in the private sector. Yet, to improve performance, measurement is essential. “If you can’t measure it you can’t manage it”. The PC has a strong Implementation and Monitoring Unit (IMU) which monitors on going and evaluate the completed projects. Also, maintains Management Information System (MIS). The IMU monitored and evaluated 732 projects, during 2007-08. Out of these 158 projects (Costing Rs. 200 million and above) problems and constraints were identified and ranked in the present study. In this connection performance management of project were measured by outputs and outcomes, client satisfaction survey, and through balanced scored approach at national, sector, institutional and project levels. The client satisfaction covers essential services, focus on processes, outputs, outcomes, and uses sample surveys. Services covered were telephones, electricity, water, health, postal, public transport, etc. The composite result indicators for effectiveness were ratio of inputs to outcomes achieved; equity indicators were measured by the extent of service by different clients groups, efficiency as ratio of inputs to outputs, while sustainability was measured by the ability to sustain operations financially. The public sector projects performance were also measured at national, sector, institutional and project level. The major problems and constraints identified were management capacity, delay in releases, concept & design problems, turn-over of PD/Staff, etc.

Are disasters any different? A new framework for impact evaluation in post-disaster settings

Presenter: Alison M. Buttenheim

Conducting rigorous impact evaluations of post-disaster aid is notoriously difficult, in part because disasters are unexpected events whose victims cannot be identified in advance. At the same time, exposure to natural disasters is rarely random, but based instead on both
 observable and unobservable characteristics of communities and households. Researchers and
development professionals therefore face three specific and related challenges to impact
evaluation of disaster relief of recovery efforts: First, the “treatment area” for post-disaster aid
is not known until after a disaster occurs. Second, an appropriate “control group” is hard to
identify. Third, several time points are relevant for impact evaluation, including pre-disaster,
immediately post-disaster, and after reconstruction. Fourth, disaster-related displacement
introduces considerable selection bias when surveying affected populations.

While qualitative case studies and analyses of convenience samples of affected areas can
provide some insight, these studies usually cannot be used to draw causal inferences about
either the specific detrimental effects of the disaster or the benefits of relief and reconstruction
aid. Ideally, impact evaluations of post-disaster aid could rely on a random sample of the
population observed prior to the disaster, which could then be tracked in the immediate post-
disaster period and throughout the reconstruction period to assess measures of health,
schooling, and household welfare. This ideal is rarely met in the field. In the absence of such
ideal samples, what other strategies can be employed to gain meaningful, rather than
misleading, insights about what relief and recovery programs work and for whom?

Using these challenges as a starting point, this paper proposes a new framework for
evaluation of post-disaster aid. This new framework is currently being developed through
reviews of existing post-disaster impact evaluations and discussions with innovative evaluators
and researchers in the disaster field; and through the careful application of evaluation “best
practices” such as experimental, quasi-experimental, and natural experimental approaches.
The framework will address the persistent evaluation challenges of appropriate counterfactual
analysis, measurement of final welfare outcomes, and sustainability. We also address the
important question of whether (and how) the impact evaluation process is truly different in the
post-disaster context.

The paper will then apply this emerging framework to the specific case of the Pakistan
earthquake of October, 2005. We rely on extensive planning and evaluation documents and
sectoral strategies from the Earthquake Reconstruction and Rehabilitation Authority and other
case studies, field reports and evaluation efforts from the post-earthquake period to ask how a
rigorous evaluation of post-disaster might have been designed differently, and, most
importantly, how disaster-prone regions can build in the capacity for post-disaster impact
evaluation in advance of disasters. We hope that this new framework will launch discussion
and productive debate about how best to measure the effectiveness of development
assistance for highly vulnerable populations in challenging post-disaster settings.

Impact evaluation and intervention to address climate change

Presenter: Martin Prowse

At the same time as climate change has reached the top of the international policy agenda,
calls for better aid effectiveness and for more evidence-based policy have led to interest in
rigorous impact evaluations (IE). This paper aims to provide and initial overview of the extent
to which IE-defined for this paper as an evaluation that focuses on the impact of an intervention on human welfare/well-being outcomes, or environmental outcomes, and which uses counterfactual analysis -could be utilized to assess climate change interventions. As the applicability to IE techniques to climate change interventions has not been widely considered, this paper does not purport to be comprehensive or exhaustive. Instead, the paper sketches out the terrain on which future studies might build. Despite these caveats, the paper offers the basis for furthering the debate on linking IE with climate change, and highlights areas where policy-oriented impact studies could be commissioned and conducted.

The paper suggests that three climate change interventions—community-based adaptation, environmental labelling and biofuel production—show considerable potential for policy orientated impact studies in the near future. This is mainly due to the medium to small size of the time and cost requirements, medium to small scale, and limited likelihood of being subject to the constraints of budget support, politics and ownership, or institutional inertia. One intervention—forest protection—shows some potential for policy-oriented studies in the near future. The major drawbacks appear to be the potential for budget support, or politics and ownership, to stymie the application of an impact evaluation methodology. Four interventions show little potential for short-term IE studies: green growth strategies; ORCHID approaches to disaster risk management; and initiatives linked to the GEF’s LDCF and SCCF. These interventions appear to be at too large a scale, and subject to the machinations of government policy and politics to such an extent that IE studies look unlikely in the short term.

Impact Evaluation in the Global Environment Facility: Looking back for Our Common Future

Presenters: David Michael Tood, Lee Alexander Risby

This paper highlights the impact evaluation design and methodological approaches developed by the Global Environment Facility Evaluation Office to address complexity, uncertainty and attribution challenges faced by Global Environment Facility projects. In doing so, we report on the testing of theory-based and quasi-experimental designs in biodiversity projects in East Africa, Central America and South-east Asia, and the selection and development combined evaluation designs for mixed-method approaches for ozone depleting substance projects in Eastern Europe, Russia and Central Asia.

The development of impact evaluation within the GEF has been driven primarily by demand from the GEF Council and other stakeholders to receive more information on results, sustainability and replication and / or scaling-up potential of positive results. However, conducting impact evaluation of GEF projects presents many challenges such as:

Complexity of the environment: Environmental change processes are ‘open-systems’ and changes (positive or negative) may only become apparent sometimes years after a project has been completed.
Reflexive complexity: GEF projects tend to the reflexively ‘mimic’ environmental complexity through their designs which often encompass a wide range of discrete and linked activities, but it is not always clear how these activities relate to higher objectives.

Data uncertainty (quantitative or qualitative): data sources on impacts are scarce in many of the GEF focal areas and there are often areas of uncertainty – no data or no agreement on existing data. Furthermore, at the project level monitoring data is often missing or incomplete.

Attribution: attribution is problematic, particularly when impacts may only be seen over long time horizon.

Given complexity and uncertainty faced by the GEF impact evaluation needs to be as much about generating insights into processes of change instigated by GEF projects as about actual demonstration of change, and indeed one cannot be seen without the other. This leads more towards assessment of the contribution and less on attribution.

We have attempted to address these constraints through the testing of a range of approaches to impact evaluation encompassing theory-based, quasi-experimental models in biodiversity and a combined theory-based and quantitative design utilizing mixed-methods in ozone depleting substance (ODS) projects. We illustrate some of the results from the case studies and also methodological strengths and limitations:

In biodiversity, a theory-based approach was tested in three GEF projects in East Africa; and a quasi-experimental approach was tested in two projects in Costa Rica and Thailand.

In ODS, a combined approach was developed based on an initial design scoping exercise, which identified availability of control groups, sources of quantitative and qualitative data. This is currently being implemented across the GEF ODS portfolio in Eastern Europe, Russia and Central Asia.

Finally, the Evaluation Office is currently developing and field testing a ‘review outcomes to impacts’ (ROTI) methodology which is based on a theory of change approach. We report on the initial experiences with this method and the possible wider evaluative application to GEF and non-GEF environmental projects where time and resource constraints prevent a comprehensive impact evaluation.

---

**Impact Evaluation of Fishery Conservation Policies in the Philippines**

**Presenter: Maria Rebecca Campos**

The Philippines is surrounded with many fishing grounds. In spite of this, most fishermen in the area live in poverty, and their plight is getting worse, not better. Current fisheries policies for the area have failed to improve the situation but no research has been done to find out why. This report uses a bioeconomic model to simulate the effects of changes in the enforcement levels of current policies. Investments of the government on different levels of enforcement were assessed using benefit cost analysis. The report assesses the effects of enforcing current fisheries policies more stringently. The situation would be transformed into...
one in which large and perhaps increasing numbers of people would continue to fish, 
expending larger amounts of effort to comply with various gear restrictions but, in all likeli hood, harvesting no fewer fish. Because the bay is already overfished, catch per unit effort and marginal productivity would decrease. Any additional fishing effort in the bay will result in a decrease in the average catch of all fishermen. Enforcement of current policies will not address the underlying problems of open access and the overfishing it leads to. One policy to deal with the problems of open access and overfishing is to set a limit on the total number of fish that can be caught and divide this quota among Lamon Bay’s fishermen.

It finds that a substantial investment would be required to ensure compliance with regulations and that the benefits of achieving high levels of compliance would exceed costs by only a tiny margin. The situation would be transformed into one in which large and perhaps increasing numbers of people would continue to fish, expending larger amounts of effort to comply with various gear restrictions but, in all likelihood, harvesting no fewer fish. Because the bay is already overfished, catch per unit effort and marginal productivity would decrease. Any additional fishing effort in the bay will result in a decrease in the average catch of all fishermen.

Enforcement of current policies will not address the underlying problems of open access and the overfishing it leads to. One policy to deal with the problems of open access and overfishing is to set a limit on the total number of fish that can be caught and divide this quota among Lamon Bay’s fishermen. Over time, the total allowable catch might be reduced. To allow flexibility, the quotas allocated to individual fishers might be tradeable. The typical Lamon Bay fisherman lies below the poverty level, has almost no secondary source of income and finds his household members willing but unable to obtain work. Efforts to reduce overfishing in the bay should therefore be complemented with measures to promote alternative sustainable livelihoods.

**Themes 2: Impact evaluation: Approaches and methods**

**Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks**

**Session number: [13.1]**

**Session stream: Environment, climate change and conservation**

**Session name: Environment, climate change and conservation**

**Evaluation and decision support for sustainable development in rural areas: case study for agri-environmental measures**

**Presenter: Jadwiga Ziolkowska**

The main objective of the paper is to present methodical approaches which can be helpful in interactive policy evaluation and design. The investigations are focused on agri-environmental measures which are obligatory instruments for policy of rural areas in all European Union (EU) member states. The objectives of agri-environmental measures which are supported both from the EU Funds and from national budgets of the EU member states include three aspects of sustainable development in rural areas: protection of biotic, abiotic and cultural resources.

National decision-makers have principally the authority to decide about financing of the measures and they steer, therefore, the directions of environmental protection and thus the directions of sustainable development. Against this background, a question arises, how to finance agri-environmental measures to achieve the objectives of environmental protection to the widest extent and which methodical approaches can be helpful to support policy evaluation and design. To answer these questions we present results of an empirical investigation for
which two methodical approaches were used: Analytic Hierarchy Process (AHP) and Linear Programming. The analysis is based on results of a case study conducted in Poland in the Subcarpathia region and on interviews with regional administration experts responsible for rural policy questions in the region. We use the Analytic Hierarchy Process (AHP) according to Saaty (1990) in order to assess the importance of agri-environmental measures ('Sustainable agriculture', 'Organic farming', 'Extensive meadow farming', 'Extensive pasture farming', 'Ground and water protection', 'Buffer zones', and 'Domestic farm animal species') with respect to the environmental objectives ('Protection of natural resources', 'Protection and conservation of biodiversity', and 'Conservation of cultural landscape') defined in the national Agri-Environmental Programme.

The results of the AHP approach were further incorporated in a Linear Programming model according to Kirschke and Jechlitschka (2002) in order to estimate objective-oriented budget allocations for agri-environmental measures from a regional perspective. The analysis provides scenarios of budget allocations for agri-environmental measures under changing economic and ecological conditions in the analysed Subcarpathia region.

The paper results prove the necessity to combine methodical approaches while solving relevant economic, environmental, and social questions. This would help to consider more comprehensively both qualitative and quantitative information and aspects of sustainable development.

Themes 2: Impact evaluation: Approaches and methods
Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks
Session number: [13.2]
Session stream: Environment, climate change and conservation
Session name: Environment, climate change and conservation

Potential and Pitfalls of Experimental Impact Evaluation: Reflections on the design and implementation of an experimental Payments for Environmental Services project in Latin America

Presenters: Jos Vaessen
Co-authors: David Todd and Gert van Hecken

The Regional Integrated Silvopastoral Approaches to Ecosystem Management Project (RISEMP) is one of the few recently completed conservation projects based on an experimental impact design, intended to allow a reliable estimate of the net effects (impacts) of the intervention. More specifically, one of the objectives of the design was to assess the impact of project incentives provided to farmers (i.e. payments for environmental services and technical assistance) on land use changes and ultimately on environmental and socio-economic benefits. This paper assesses the merits and limitations of the experimental design as a basis for impact evaluation of the project. The following questions are addressed: 1) What are the merits and limitations of the experimental design in supporting claims on impact given the project’s objectives? 2) To what extent is the validity of the design affected by the nature of the design and implementation process in the field? 3) On the basis of the foregoing, what is the utility of the experimental design in supporting claims on the project’s impact? 4) What are useful complementary and alternative methods for assessing the project’s impact? The analysis mainly focuses on one of the project’s three pilot sires; in Nicaragua. It shows that the utility of the experimental design to determine attribution was compromised by selection bias, contagion and unintended behavioral responses. The paper illustrates how the analytical
potential of experimental impact evaluation must be supported by careful attention to specific features of the design and implementation of the experiment. It also argues that the combination of theory-based evaluation with an experimental design approach can maximize the precision and explanatory power of impact evaluation.

<table>
<thead>
<tr>
<th>Themes 2: Impact evaluation: Approaches and methods</th>
<th>Session number: [13.2]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks</td>
<td></td>
</tr>
<tr>
<td>Session stream: Environment, climate change and conservation</td>
<td>Session name: Environment, climate change and conservation 2</td>
</tr>
</tbody>
</table>

Ecotourism and the Sustainable Development of protected areas: the Index of Sustainable Ecotourism Impacts. The case of Wadi El Gemal National Park in Egypt

**Presenter: Cherine Khallaf**

Ecotourism is a concept defined by its positive impacts that encompass conservation and the sustainable development of the destination. However, acute deficiencies in the identification and measurement of the activity’s comprehensive impacts were revealed in the literature, thus, leaving the concept void of content. This research developed a general framework for ecotourism sustainability in protected areas that was made operational through a conceptual approach applied to the Indicator of Sustainable Ecotourism Impacts (ISEI). The index use and limitations were illustrated in the case of Wadi El Gemal National Park in Egypt. This paper argues that the complex ecotourism impacts, describes in the Protected Area Ecotourism Sustainability Framework (PAESF), can only be captured through inter-disciplinary assessment approaches. The Index of Sustainable Ecotourism Impacts was therefore constructed following the model of the alternative measures of welfare. ISEI was further illustrated through its application to the case of Wadi El Gemal National Park in Egypt. The results highlighted the sustainability risks and opportunities present in the case studied, while illustrating how ISEI application was strongly hindered by data deficiencies and lack of integration of social, economic and ecological statistics at the local level, thus, revealing an important deficit in ecotourism systems in developing countries such as Egypt.
What are the Effects of Land Use Restrictions on Local Communities? Evidence from an Impact Evaluation of Costa Rica’s Protected Areas

**Presenter:** Kwaw S. Andam

**Introduction**

Global efforts to reduce deforestation in developing countries rely heavily on protected areas and land use restrictions. The social impacts of protected areas on local communities are currently the subject of heated debate among conservation and development experts. Measuring these impacts is difficult because the effects cannot be isolated from other factors.

We address this difficulty by applying a quasi-experimental approach (matching methods) to measure the impacts of pre-1980 protected areas in Costa Rica on socioeconomic outcomes in 2000. We address three methodological issues: (1) selection on observables: overt bias may lead to incorrect impact estimates, especially given non-random placement of protection; (2) spatial spillovers: protection may affect outcomes in neighboring unprotected communities; and (3) hidden bias: unobservable differences between protected and unprotected lands may confound impact assessment.

**Data.** Our socioeconomic data are from censuses in 1973 and 2000 (Instituto Nacional de Estadística y Censos) and GIS data are from the University of Alberta, Canada, the Instituto Tecnologico de Costa Rica, and the Instituto Geográfico Nacional. The unit of analysis is the census segment. Our sample comprises 4,694 census segments surveyed in 1973. We measure socioeconomic outcomes using a poverty index (Cavatassi et al. 2004), and by percentage of houses (a) in bad condition; (b) in slum areas; (c) without a telephone; (d) without electricity; (d) without water supply.

**Methods.** The key challenge is to control for biophysical and socio-economic covariates that affect both changes in social welfare and the location of protected areas. We establish this control through matching methods, which are increasingly used to test cause-effect relationships using non-experimental data. Matching works by identifying, ex post, a comparison group that is “very similar” to the treatment group with only one key difference: the comparison group did not participate in the program (Rubin, 1980; Rosenbaum and Rubin, 1983; Imbens, 2004).

We match communities based on forest area in 1960, land productivity, “road-less volume” (Watts et al., 2007), distance to major cities (markets), and baseline socioeconomic conditions. Based on covariate balance (Ho et al., 2007), we use Mahalanobis distance matching. We also estimate the spatial spillover effects of protection on unprotected segments near protected areas, and we test for the effect of potential hidden bias on our estimates.

**Findings.** We find that Costa Rica’s protected areas had positive socio-economic impacts. Communities affected by protection had a lower poverty index, and protection led to better outcomes in housing, living conditions, and access to electricity. These findings are different from findings with conventional empirical methods, suggesting that failure to control for confounding factors or baselines can lead to inaccurate estimates.
Do the Poor Benefit from Devolution Policies? Evidence from Malawi’s Forest Co-Management Program

Presenters: Charles B.L. Jumbe, Arild Angelsen

Do local people, and in particular vulnerable groups, benefit from devolution of forest management? This question is analyzed by using the propensity score matching and decomposition techniques on household data from Chimaliro and Liwonde forest reserves in Malawi. Despite contrasting evidence of the impact of co-management between the two sites and the fact that program participants have substantially lower income than non-participants, overall results indicate that the livelihood of participants would have worsened in absence of the program. We find that forest co-management increases monthly revenue for participants by between MK17 (US$0.22) and MK18.00 (US$0.24) per household, an increase of between 37% and 41%, while both female-headed and low-income participants earn extra income of between MK14 (US$0.18) and MK29 (US$0.39) per month per household, an increase of between 22% and 51%. However, decomposition results suggest that male and high-income participants capture most benefits from co-management due to statistical ‘discrimination’ and productivity differences which account for 100% and 60% of the inter-group income differential, respectively.
Potentials and challenges of using M&E for institutionalizing impact evaluation and adaptive learning within organizations and programs involved in climate change adaptation

*Roundtable to be organized by the Climate Change adaptation in Africa Program (CCAA), a joint initiative between the UK department for international development (DFID) and the International development research centre (IDRC)*

The CCAA program has been paying a lot of attention to the practice of monitoring and evaluation (M&E). In deed, the program needs to demonstrate its outcomes and its contributions to increasing adaptive capacity of African people and organizations to climate change. In addition to this, the practice of M&E itself can greatly increase learning from experience. Because adaptation is a learning process, involving anticipating future changes, paying attention to current changes, trying out adaptation strategies and verifying how well they work, M&E can be used to strengthen adaptive capacity. The CCAA program has been encouraging the use of M&E within the projects its supports. More recently, it has partnered with some regional and sub-regional organizations aiming at institutionalizing the use of M&E in strengthening adaptive capacity to climate change. In deed, such organizations are already conducting M&E, often in collaboration with national partners who, in turn, collaborate with local partners.

However, there are many challenges for M&E to contribute to impact assessment, learning and adaptation. There is often a lack of communication between organizations and between different administrative levels. Although each level needs the others to produce and to evaluate impact on the environment and on ultimate beneficiaries, it is very difficult for the distinct levels to obtain and to use each other’s data. Often there is a lack of human, technical and financial resources in the organizations that are tasked with M&E. But even when all the resources are there, there is often an incompatibility between methods and indicators used. M&E is often not well integrated into planning and decision-making processes. Also, M&E often concentrates on biophysical and socio-economic conditions to be affected and on activities and outputs of the program being monitored without giving enough attention to the multiple contributions of a large number of organizations involved in creating the desired impact. To address these gaps, a number of regional and sub-regional organizations have been developing an initiative with CCAA to strengthen their capacity to work with national and local organizations in conducting M&E of- and for- adaptation. It is anticipated that partnerships between administrative levels could be catalyzed by approaches that recognize changes in organizational practices of partners as outcomes of policies and programs.
Rigor and ethics in impact evaluations in Africa

Presenter: Donna M. Mertens

Co-authors: Bagele Chilisa, Tony Onwuegbuzie

This panel will address the challenges in conducting rigorous impact studies in Africa and other low-income regions from an ethical perspective. The relationship between ethics and rigor is at the same time obvious and complicated. The obvious part is that ineffective programs waste time and money and contribute not only to poorer quality of life, but also to potential loss of life. The complicated part rests in the legacy of colonialism in Africa and other low-income regions and the consequent power differentials that influence access to resources and quality of life.

Mertens will examine the challenges associated with impact evaluation for social transformation in the post-colonial and indigenous context, including issues of power that arise from that historical legacy, as well as from current societal structures that reflect on-going inequities. How can reflective stances be developed that address these issues of power differences, discrimination, and oppression that are ethically defensible? What are the philosophical assumptions that provide guidance in these culturally complex contexts? The axiological assumption prioritizes social justice and human rights and thus provides grounding for consequent ontological, epistemological, and methodological assumptions for evaluation that is specifically targeted for social transformation amongst post-colonial and indigenous peoples. A combination of contributions from the transformative paradigm and the scholarship of postcolonial and indigenous peoples offers potential pathways to ethical practice in the use of evaluation for social transformation purposes. Examples are drawn from the Handbook of Social Evaluation Ethics (Mertens & Ginsberg, editors, 2009, Sage Publications) who represent indigenous peoples from Africa, North America, and the Australasian area, as well as Western scholars who wrestle with these crucially important issues in terms of planning, conducting, and making use of evaluation that is designed to further human rights and social justice.

Chilisa will examine the challenges that postcolonial/indigenous scholars present to methodological imperialism and will discuss a framework to discuss indigenous-based methodologies based on ethics and rigor in impact evaluation. She will explore the meaning of postcolonial/indigenous evaluation methodologies and the ontological and epistemological assumptions in postcolonial indigenous evaluation methodologies. The argument is that postcolonial indigenous methodologies are founded on a critique of a historical ontology emanating from colonialism and imperialism. The critique uses postcolonial theory, critical theory, and critical race theory in particular as tools of analysis and gives rise to indigenous evaluation methods of critique and resistance. Postcolonial/indigenous evaluation methodologies further draw from indigenous knowledges to explain new forms of methods for evaluating impacts that are informed by a relational ontology and epistemology. In the indigenous evaluation paradigm, evaluation is viewed as a relational process. Postcolonial/indigenous societies for instance, perceive human beings as directly connected with the entire universe, spiritually and materially. The emphasis on connectedness points to wholism as an essential concept that should inform ways of knowing and the values that guide the evaluation process. A relational epistemology draws our attention to relational forms of knowing as opposed to individual descriptions which according to Thayer-Bacon (2003) have dominated Euro-Western theories on ways of knowing for a long time. She will illustrate
methods of resistance and critique that draw from indigenous knowledges, world views and assumptions about a relational ontology and epistemology.

Onwuegbuzie will examine the issues and challenges in the African context involved in interpreting, reporting, and disseminating findings from rigorous impact evaluations that utilize a mixed methods framework—namely mixed methods impact evaluations (MMIEs). In terms of interpreting MMIE findings, 11 mixed methods evaluation paradigms/worldviews will be compared and contrasted with respect to three axiomatic components (i.e., ontological, epistemological, and methodological foundations), and seven issues (i.e., nature of knowledge, knowledge accumulation, goodness or quality criteria, values, ethics, inquirer posture, and training). Based on their philosophical underpinnings, each of these mixed methods evaluation paradigms/worldviews will be classified according to their potential for leading to strong meta-inferences (i.e., inferences stemming from both the qualitative and quantitative findings that are combined into a coherent whole; Tashakkori & Teddlie, 1998). Also, Onwuegbuzie and Johnson’s (2006) nine-component typology will be used to outline how the legitimation of meta-inferences in MMIEs can be assessed. With regard to reporting MMIE findings, Onwuegbuzie and Leech’s (2004, 2005) typologies will be used to provide a framework for reporting “significant” findings stemming from the quantitative and qualitative phases of MMIEs. Finally, the five authenticity criteria developed by Guba and Lincoln (1989) (i.e., fairness, ontological authenticity, educative authenticity, catalytic authenticity, and tactical authenticity) will be used to provide guidelines for disseminating findings stemming from MMIEs. Throughout this presentation, Onwuegbuzie will make clear the importance of rigor and ethics in impact evaluations in Africa and beyond.

Themes 2: Impact evaluation: Approaches and methods
Session number: [15]
Session stream: Evaluating aid
Session name: Evaluating aid

The Comparative Method for Assessing Impacts and Cost Effectiveness of Development Assistance Programs

Presenter: Paul Clements

This paper argues that official development assistance programs should routinely be evaluated in terms of their impacts and cost effectiveness based on the comparative method. Today’s most common approach is based on the DAC criteria: relevance, effectiveness, efficiency, impact and sustainability. Even though “impact” is one of the DAC criteria, however, these evaluations seldom make sustained efforts to quantify total impacts. Usually the question is interpreted generally, as, for example, “did the intervention have positive impacts?”

Evaluations should help the development community to reduce world poverty as much as possible with available resources. To provide a basis for learning and accountability to this end, evaluations should routinely estimate (1) total impacts attributable to the intervention, understood as total changes in beneficiary conditions (past and future, intended and unintended) compared to the situation one would expect without the intervention, and (2) the intervention’s cost effectiveness, understood as the magnitude of total impacts in relation to total costs. This can be achieved through the comparative method:

Proceed on a sectoral basis, starting with water, primary health care, and microfinance projects. Synthesize existing evaluations and relevant studies, identifying factors that contribute to lesser or greater impacts and the main counterfactual “threats,” and identify practical and fairly reliable means for assessing levels of and changes in the magnitudes and relevant qualitative conditions of these factors. These might include background investigations and sample surveys. Find and/or carry out several careful evaluations, making
coherent efforts to estimate impacts and costs. Assign values to qualitative and non-economic impacts based on the donor’s willingness to pay and weight impacts based on need. Estimate each project’s cost effectiveness like an economic rate of return, identifying the range of plausible values (i.e. where higher and lower values can confidently be excluded). This yields a distribution of projects based on their relative cost effectiveness.

Get several evaluators to train themselves on the identified means for estimating impacts and cost effectiveness in the selected sectors and to familiarize themselves with the strong evaluations and relevant studies. Secure agreements with donors for these evaluators to evaluate their interventions in the selected sectors. The evaluators locate each new project in the established distribution, noting upper and lower boundaries of plausible impacts, and checking to confirm that procedures and coefficients are consistent between evaluations. Form a review board that can expel an evaluator who diverges too greatly from the established approach.

This will yield:

- estimates of impacts and cost effectiveness that are reliable within a reasonable range,
- better explanations of how design and implementation affect cost effectiveness,
- stronger assessments of strategic positions, and
- better recommendations.

It will support better learning and accountability for donors and across the development community and improve the incentive environment facing planners and managers.

---

**Themes 2: Impact evaluation: Approaches and methods**

**Session stream: Evaluating aid**

**Session number: [15]**

**Session name: Evaluating aid**

**Measuring and Reporting Results: Experiences of the Donor Committee for Enterprise Development (DCED)**

**Presenter: Jim Tanburn**

There is little information about the results of development initiatives, that is quantified, comparable and credible; the Donor Committee for Enterprise Development considers that this is at least partly because impact evaluation has often been seen as an event, imposed on programme managers from outside, and therefore not gaining their full collaboration. This is a critical constraint, in a field that is so technically challenging; meanwhile, many programme managers in the field are fiercely committed to being effective – and would like a channel through which to communicate their achievements.

The DCED has therefore worked with a range of donor-funded programmes during 2008 to identify the essential elements for a results measurement system: clear articulation of the results chains, definition of the right indicators, good measurement practices, attribution of the measured changes, capturing wider changes in the system, and relating results to programme costs. Credibility is assured by an external auditor, who examines the measurement process, to determine whether it meets minimum standards. Comparability is assured by universal adoption of three impact indicators: scale, changes in income, and changes in employment.

Work during 2008 has shown that programme managers in the field are willing and able to participate in such a system – although some support and training are needed to bring staff up to speed. During 2009, the system will be expanded to include more programmes, particularly in Africa and Latin America. The DCED will also work to mesh the system more with existing
results measurement practices; for example, could independent evaluation departments play the role of auditor, examining the results measurement processes? Can programmes generate enough data, using this system, for independent evaluators to use – for example to map enterprise-level impacts onto household-level impacts on poverty? Can the approach be applied by partner governments?

A training course will also be developed, to include all of the key skills required; supporting infrastructure, particularly to include a cadre of suitable consultants, will also be further expanded. Experience suggests that managers need help in particular to articulate the results chains of their programmes more explicitly. Empirically, one has to conclude that logframes have not focused minds on results measurement; the logframe format is too crude to represent the complex results chains implicit in large, donor-funded programmes.

Further information can be found on the Donor Committee website, at www.Enterprise-Development.org/page/Measuring-and-Reporting-Results There is also much experience that can be reported on in Cairo, for example around the ‘mock audits’ of five donor-funded programmes on private sector development (PSD) in Asia during 2008. There will also be an opportunity for colleagues in the Cairo Conference to join the DCED process, going forward; collaboration is particularly sought both from evaluation professionals, and from programme managers. At the time of writing, the Annual Meeting of the DCED has just strongly endorsed the progress in 2008, and the plans for 2009; the DCED thematic steering group, including *inter alia* UNIDO, GTZ, ILO, EC and the Government of the Netherlands, is planning to hold its next meeting around the Cairo Conference.

**Themes 2: Impact evaluation: Approaches and methods**

**Session number: [15]**  
**Session stream: Evaluating aid**  
**Session name: Evaluating aid**

**Development per Aid Dollar - A Data Envelopment Approach**

**Presenter: Bo Sjo**

This evaluation applies a value-added Data Envelopment Analysis (DEA) to evaluate the impact on aid on development. The advantages of the DEA approach are that it allows for development to be measured as a vector of outcomes, instead of focusing only on economic growth. The method is focused on measuring relative efficiency and is not dependent on specification of ‘production’ that clearly relates aid as an input to development outcomes. The disadvantage, in this context, is that the choice of outcomes is given by available data, and without the backup by a clear theory the results can hard to understand. Further, the DEA approach is limited in terms of the number of inputs and outputs that can be used for the 60 countries in the sample.

In this evaluation, development outcome is represented by changes in PPP-adjusted GDP, increase in school enrollment, decrease in under-five mortality rate and increase in civil and human rights as reflected in the Voice and Accountability index. In the first run the input is cumulative ODA received in previous period. The results show that the more aid a country has received the less it has developed. This is so for all donors as well as for Sweden.

These results reflect that aid in not a constraint for development. Countries that develop do so without aid. The results might also reflect that donor countries are super-selective in their allocation of aid. In anticipation of the outcome donor countries allocate more aid to countries with less expected development. Thereby the effect of aid is to reduce the distance between those countries that develop with little aid and those that do not. The set-up of the DEA in this
evaluations do not allow for a test of the latter hypothesis, since that would require a more theory based approach.

Recognising that aid is a highly selective process, a second DEA is set up where the input the sum of ODA in relation to the initial level of Voice and Accountability index and the level of GDP. This breaks the negative correlation between aid and development. In terms of relative development conditional on the amount of aid given selectively, countries such as Algeria, Burundi, and Cambodia are effective. While Bolivia, Papua New Guinea and El Salvador are the least effective.

<table>
<thead>
<tr>
<th>Themes 5: Capacity building for impact evaluation</th>
<th>Session number: [16 ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session stream: Evaluating capacity building interventions</td>
<td>Session name: Evaluating capacity building interventions</td>
</tr>
</tbody>
</table>

**Case Study: Challenges of Measuring Impact Evaluation of Capacity Building Project in Bukoba rural, Republic of Tanzania**

**Presenter: Neeta Singh**

**Co-authors: Dorothy Ettling, Paula Caffer, Alison Buck, Susana Deku**

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 to these researchers is the impact that our 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. Though women are de facto heads of 25% of the total households, their average income is 45% below that of man-headed households; and an estimated 69% of these households live below the poverty line. 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.

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). The difficulty in assessing impact hinders the potential for further funding for program continuity and expansion. 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. On-going evaluation since 2004 has evidenced changes in women’s attitudes, behaviors and family and community involvement. However, it is still complex to mention true impact. For example, a woman will indicate in an outcome evaluation, “This project enables me to pay school fees for my daughter and buy beans for our table.” To convert this important information into existing economic indicators, thus satisfying stakeholders, is problematical.

To address this challenge, 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.
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.

**Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks**

**Session number: [17]**

**Session stream: Evaluating partnerships**

**Session name: Evaluating partnerships**

**Evaluating Partnerships for Sustainable Development: Case Study of PSRHH**

*Presenter: Mark Ojukwu*

Partnership in the recent times has been identified as a tool by which to achieve development outcomes in poor resource settings. Partnership however is a complex concept which tends to respond to address issues by ensuring the achievement of sustainable development outcomes. Partnerships between agencies, professions, and across sectors in the delivery of public services are now a major policy goal across the globe - partnership globally is the new language of public governance. The emergence of partnership working therefore reflects a desire to move from the development of policies and programmes as well as the delivery of services within fragmented organizational and professional microcosmic blocks to an integrated multi-sectoral and multi-professional approach based on comparative competence to deliver on sustainable development outcomes. Be that as it may the importance of structured reflection, learning and evaluation of partnerships such as the PSRHH programme cannot be under emphasized.

PSRHH partnership evaluation is used to increase accountability; promote learning and bring about programme improvements. The evaluation of the partnership thus includes evaluation of partnership activities (assessing the operations and impact of the partnership driving the PSRHH programme) and the partnership working (measuring what differences the partnership is making towards delivering sustainable development outcomes). PSRHH partnership evaluation approaches are underpinned by a set of theoretical assumptions based on the following parameters: Value-for Money Analysis; Outcome-Focused Evaluation; Process-Outcome Evaluation; Stakeholders’ Analysis; and Evaluation of the Partnership Mechanism. The PSRHH partnerships evaluation approach thus highlighted the partnership’s success factors vis-à-vis its impact; vision and leadership; partnership dynamics; strategy and performance management; influencing; participation and cost-effectiveness adopting an ongoing participatory and formative evaluation of the partnership process.

In order to ensure that the principles and operating procedures adopted by the partnership are being followed, the PSRHH partnership conducts an ongoing participatory evaluation involving all the implementing partners. The PSRHH partnership evaluation process follows a vicious circle formed by evaluation, learning, improvement, measurement and evaluation. The facilitation of the PSRHH partnership evaluation requires analysis of the evidence captured, diagnosis of partnership strengths and weaknesses, and action planning for improvement. The resulting evaluation report are used to demonstrate to partners the value of their participation as well as the achievement of the partnership by ensuring that the partnership focuses on priority objectives and challenge poor performance. Improve decision making by providing feedback on progress, identifying areas where action is required to improve performance and reviewing resource allocation; and provide the basis for learning and development. As this occurs, documentation of the partnership evaluation decisions helps in creating mutual understanding and addressing issues capable of derailing the partnership.
Partnership evaluation often neglected is required to allow judgment to be made about the overall effectiveness of any partnership, facilitate debate about the sustainability of the partnership and its intended outcomes, as well as the future direction of the partnership. Besides partnership evaluation is expedient as it identifies tangible areas for improvement, and learning.

**Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks**

**Session number: [17]**

**Session stream: Evaluating partnerships**

**Session name: Evaluating partnerships**

**Evaluation of the New Partnership for Africa’s Development: Approach and Methodology?**

**Presenter: Alain A Ndedi**

Many African states at independence wanted to be popular with the African masses. They, as a result, came up with many development projects than what the local resources could finance. The option was to supplement the local capital base with foreign capital. In the more recent past, the cold war and post cold war politics, protracted conflicts, structural adjustment programmes (SAPs) and HIV/Aids pandemic have left the continent poorer than before. Unlike East Asia, which has enjoyed a dramatic reduction in the number of people living in poverty over the past 15 years, sub-Saharan Africa has seen dramatic increases in both the total number of poor people and the fraction of its population that is poor (World bank, 2000).

In order to deal with the above mentioned challenges, various initiatives have been adopted even though most of them are tantamount to rhetoric. From Tony Blair’s Africa commission, the G7 Finance Ministers debt relief, the live 8 concerts, the make poverty history campaign and the G8 Gleangles promises to the United Nations 2005 summit, Africa’s gains seem to be reduced to some kind of public relations exercise indicating that the world has taken stock of the plight of the continent. However, since 2001, the New Partnership for Africa's Development (NEPAD), a vision and strategic framework for Africa’s renewal was launched as a driver for African countries to move from long severe poverty to sustainable economic development. Almost seven years after the launch of the initiative, African states in general are still suffering from the same ills.

The current contribution is an attempt to propose an approach for the evaluation of the NEPAD programme in order to effectively and efficiently deal with real challenges that the continent is facing. The first part of the paper presents a glimpse of the current continental economic/political/social situation and the need for immediate actions. The second part defines the concepts of NEPAD and evaluation and the importance of the latter in the development agenda of the African continent. The third part proposes a methodology for evaluating the NEPAD programme with specific reference to the redefinition of its strategic framework in a context of a more globalise and competitive world.
Evaluating the impact of a research-grant programme

**Presenter: Kevin Kelly**

A conceptual approach and method for evaluating the impact of a research-grant programme is presented. The five-year funding programme supported research on the social impacts of HIV/AIDS in South Africa. The ultimate aim of the funding programme was to stimulate research leading to more evidence-informed policy and practice. It proved difficult to conceive a model for assessing impact in this context, given the weak existing links between policy and practice, and the poor orientation to evidence in the policy environment. An alternative, outcomes-oriented evaluation model was developed, on the basis of which ‘contribution to impact’ could be inferred. This was based on a ‘critical realism’ approach to evaluation. The starting point was a review of how policy and practice in this context might be theorised to change, as a result of the kinds of processes supported and outputs delivered by this programme. This led to an understanding of ‘pathways’ to research impact, as a result of which changes in policy and practice might reasonably be expected to take place. This ‘theory’ or model of change was further refined in early interviews with programme grantees. The model maps the range of ‘contributions to impact’. An interlinked group of outcome indicators was developed. These mostly consisted of rating-scales to quantify the degree of achievement on the part of each of the grantees, for each of the areas of contribution to impact. The programme was critically evaluated in each of these areas. Rather than presenting the findings of the evaluation, the presentation focuses on: the challenges of assessing the impact of research processes aimed at bringing about changes in policy and practice; and findings related to what needs to be done to optimise research-grant making in this and similar contexts.

Building Capacity on Program Evaluation in Latin America: The Experience of the Partnership between Mexico’s National Institute of Public Health and the MEASURE Evaluation project

**Presenter: Gustavo Angeles**

**Co-authors: Jose Urquieta-Salomón, Bernardo Hernandez-Prado; Mara Tellez-Rojo, Aurelio Cruz-Valdez**

In Latin America, as in many other developing regions, there is a shortage of professionals with program evaluation skills. That shortage is particularly severe in government agencies and NGOs implementing population, health and nutrition programs. Unfortunately, there are few places in Latin America where professionals can go to increase their skills in program evaluation. To address this gap, since 2004, the National Institute of Public Health (INSP) of Mexico and the USAID-funded MEASURE Evaluation project entered a partnership for establishing training programs in evaluation and strengthening capacity to carry out impact evaluation studies and to
provide technical assistance. The partnership included jointly conducting annual workshops on program impact evaluation, supporting the development of a track on program evaluation as part of the Masters of Health Sciences and Masters of Public Health, training of key faculty on advanced methods, conducting collaborative evaluation studies of the impact of Mexican government health programs, transfer of materials, provision of fellowships, and short- and long-term exchanges of faculty for impact evaluation studies. Additional activities have been implemented for advancing technical and financial sustainability. The impact evaluation training programs have been highly demanded and successful. The number of applicants for the workshop increased from 42 in 2004 to 107 in 2008; in total 85 professionals from 16 countries have been trained on impact evaluation. Many participants report the application of learned skills to their work. The program evaluation tracks have been established in the two Masters programs and over 30 students have graduated or are pursuing the diploma in M&E. In terms of institutional development, the partnership contributed to the already ongoing process of INSP defining its role on evaluation in the public health field. INSP was in the process of assigning higher prominence to program evaluation in its teaching, research and practice areas. In recent years, INSP has adopted “Health program and policy evaluation” as a Strategic Line of Action, it has created a “Center for Surveys and Program Evaluation”, and has taken leadership roles in multiple impact evaluations.

This experience provides some lessons on capacity building partnerships: It requires long-term commitment from both parties; identify “Champions” in M&E within the institution; the institution identifies a “market for program evaluation” and is able to respond to it; the institution identifies program evaluation as a strategic area for development; the institution facilitates provision of incentives (either economic or professional advancement) for faculty participation in M&E activities; and it is necessary to address technical, financial, and administrative capacity to achieve sustainability.

### Themes 5: Capacity building for impact evaluation

<table>
<thead>
<tr>
<th>Session stream: Evaluating capacity development</th>
<th>Session number: [19.1]</th>
</tr>
</thead>
</table>

Towards Becoming an Authority in Impact Measurement and Knowledge Management: Building capacity in applying DME standards in CARE Sierra Leone

**Presenter:** Bockarie Sesay

**Co-authors:** Ahmed Ag Aboubacrine, Chris Necker

Over the past few years, CARE International has developed a series of guidelines, standards and materials on design, monitoring and evaluation (D,M&E) in order to improve the quality of programmes and projects worldwide. The CARE International Sierra Leone Strategic Plan for 2007 – 2011 includes the following enabling strategy: “Strengthen CO and partner organizational capacity to implement accountable, high quality programming driven by thematic expertise and robust DME, including knowledge management and reflective practice”.

The first key activity of this enabling strategy is: CARE and partners are implementing effective DME practices, including reflective practice and knowledge management. One measure of that activity is the recognition of CARE Sierra Leone as an authority in impact measurement and knowledge management, especially in the areas of Household Livelihood Security (HLS), health, youth and governance. In doing so, a DME capacity assessment has been undertaken by a task force.
The purpose of this round table is to present the methodology used in Sierra Leone to tackle the challenges of building staff and partners (local NGOs, line ministries, key stakeholders) capacity in order to create an impact-led culture (thinking evaluatively, Rugh 2006) through the systematization of norms and standards simultaneously with building technical skills while promoting adaptive solutions.

The methodology encompassed a DM&E capacity assessment for individuals that led the establishment of a coherent capacity building plan. Alongside that assessment, another assessment target the status of adaptive challenges related to program quality at both local and national level. The later assessment covered areas such as program coordination, knowledge management and learning, communication / dissemination, documentation, quality assurance structure, and country office management information system.

The findings have been analyzed and led to a reform in CARE Sierra Leone structures and systems that would enable the achievement of the above goal.

The capacity building targets program staff comprises CARE Sierra Leone DME staff, Managers, Officers, Supervisors, Advisors as well as staff from partners through a combination of ad-hoc trainings, on-the-job coaching and special mentoring by external resource person including consultants. Implementing this capacity building plan needed financial resources as well technical coaching and mentoring delivered by various actors (Senior staff, Universities, Private firms, Free lance consultants).

It is noticed that the needs were huge due the decade of civil war which provoked a brain drain from Sierra Leone to UK and US. Therefore the capacity building plan took into account only the issues which are relevant for the current projects and at the same time achievable within two-year time. So, a prioritization was done based on the most common needs among staff and that led to a list of specific capacity building events and processes.

This article present the status of the above events and processes and the extent to which their implementation is foster or not CARE Sierra Leone capacity to do effective program and project diagnosis and design, establish and implement useful monitoring systems, and organize good quality evaluations.

### Themes 5: Capacity building for impact evaluation

**Session stream: Evaluating capacity development**

**Session number:** [19.2 ]

**Session name:** Evaluating capacity development

---

**Evaluation Capacity Building in South Asia: Experiences, Lessons Learned and Way Forward**

**Presenter: R.S. Goyal**

In South Asia region there is a dearth of professional evaluation expertise in almost all areas. A situational analysis indicates that evaluation professionals lack access to high quality professional development training in evaluation, which consequently has hindered the evaluation capacity in the region. Evaluations tend to be put of due to weakness in program design and implementation. More broadly, lack of emphasis on the role of the evaluation within an organization’s culture of learning and accountability has also limited evaluation capabilities. The capacity building is required at all levels; regional, national and sub-national.

This paper shares experiences in capacity building in outcome and impact evaluation of program professionals in development sector.
Imputes for this initiative was provided by Unicef ROSA by inviting institutions from South Asia to join in a capacity building program in evaluation. After initial rounds of discussion between the institution in South Asia and professional training institutes from the west, a Training of Trainers program was evolved and delivered on December, 2007 under the author’s leadership. This program was attended by professional trainers from India, Pakistan, Bangladesh and Sri Lanka. The training broadly covered the themes of; types of evaluation, evaluation design, ethics, best practices, writing terms of reference and others. Feedback of this one week training program was critically examined with reference to contents, training pedagogy, background and reading material provided. The result shows that the mode of delivery of training was good. However, contents need modification to suite to the background and needs of the participating professionals and the programs in their respective countries.

As a next step, the author (other participating institutes as well) has organized a training specific to evaluation professionals in health sector on 13-17 October, 2008 at Jaipur, India. This training has focused on; developing understanding on utilization of evaluation findings for program management, result based management, types of evaluation and links to monitoring, evaluation design, ethics and standard of evaluation, developing terms of reference and process of carrying out evaluation. The feedback of this evaluation training shows to it has been closed to the expectation of the participants and facilitated the enhancement of their skill in the sphere of planning, carrying out and utilization of evaluation finding. It is proposed that the participants of this training will be re-visited after the gap of one year to access usefulness of training in their working.

Using the lessons learned from these training programs, few more evaluation training are planned. The training module will also be standardized for wider use in the region and elsewhere. A one year virtual diploma program in evaluation for the development professionals is also being developed in the context of South Asia and will be rolled out in Sept./Oct.08.

| Themes 5: Capacity building for impact evaluation | Session number: [19.2] |
| Session stream: Evaluating capacity development | Session name: Evaluating capacity development |

A Model for Using Technical Assistance to Improve the Evaluation Capacity of Local Programs

**Presenter: Wayne M. Harding**

**Co-author: Cheryl Vince-Whitman**

For over a decade, federal and state agencies and foundations in the U.S. have promoted the use of prevention programs that rigorous evaluations have shown achieve the impacts they target, and many of these organizations now fund only such proven programs. The supply of these programs is so limited, however, that often it isn’t possible to identify one that fits local needs and conditions. In response to this problem, since 2004, the federal Center for Substance Abuse Prevention has funded Service to Science (StoS). This initiative is based on the view that many local programs do prevent substance abuse, but there isn’t good evidence to show they do, and without this they will not, and probably should not, be widely disseminated. STS is designed to: help innovative, local prevention programs document evidence of their effectiveness, and increase the number of effective programs from which communities and States can select to prevent substance abuse and related health problems.
As of June 2008, STS had provided free technical assistance (TA) to 239 prevention programs. Also, since June 2006, 79 of these programs, after receiving approximately one year of TA, have successfully competed for one-year STS contracts of up to $30,000 to continue to build their evaluation capacity. Surveys and case studies indicate that STS has succeeded in serving a wide variety of programs. For example, a 2007 survey of 93 programs shows that they addressed alcohol abuse (80%), illicit drugs (60%), tobacco (58%), violence (37%) and/or HIV/AIDS (9%). And, while most programs served adolescents (71%) or children ages 6-12 (51%), many served populations for which there are few evidence-based programs: 19% served young adults 18-25 and approximately 10% served children from ages 1-5, adults 26-55, and/or adults over 55. The same survey indicates that the TA was valued and successful. Asked to rate STS, 72% of respondents were satisfied with it, 91% found it useful, and 91% indicated they would recommend it to others. And, when they were asked in an open-ended item to identify factors that facilitated their program’s progress toward becoming more evidence-based, their most common response (79%) was the TA received from STS.

This paper identifies key characteristics of STS associated with its success. These include: using TA providers with experience evaluating the same types of programs they assist, carefully matching TA providers to the individual program’s needs, providing long-term TA that typically lasts one year, and building a strong personal relationship with programs through a STS TA “case manager” who orchestrates services for the program.

STS is cost efficient. Rather than funding both the development and evaluation of new prevention programs, STS assists programs that are already operating and have made some evaluation efforts. These programs can often make substantial gains with relatively few additional resources. The judicious application of lessons learned from implementing STS should be useful for building evaluation capacity in a variety of settings/cultures. It has already been applied, with some adaptations, to prevention interventions that serve the Pacific Jurisdictions, Native Alaskans, and American Indians.

<table>
<thead>
<tr>
<th>Themes 5: Capacity building for impact evaluation</th>
<th>Session number: [19.2]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session stream: Evaluating capacity development</td>
<td>Session name: Evaluating capacity development</td>
</tr>
</tbody>
</table>

**Increasing the capacity of MOH to use the data for decision-making**

**Presenter: Anna Blyumina**

With funding from USAID and the support (financial, logistical, legal, and technical) of the MOH and participating facilities, JSI Maternal and Infant Health Project provides a full cycle of effective evidence-based perinatal practices in more than half of regions of Ukraine covering currently 120 maternities. The Project assists the Ministry of Health of Ukraine with reviewing and revision of existing protocols and also develops new EBM standards for mother and newborn health care, with a stronger focus on integration of these standards into a package of EBM perinatal practices in Ukraine. To ensure the sustainability of the results, the Project revises current curricula for pre- and post-service medical institutions (medical universities and colleges) in order to increase the evidence base of educational programs for medical students and health care providers including OB-GYNs, midwives, and nurses. In cooperation with WHO European Division, the Project has developed an effective evidence-based training package on antenatal, perinatal, and pediatric care.

MHIP has established a sound monitoring and evaluation (M&E) system that cover all the maternities working with the Project. The system measure, analyze, interpreters and report on routine maternal and newborn’s health care procedures and results in a way that provides accurate
and useful information for improving in maternal and newborn health care. One of the main purposes of the MIHP M&E system is to trace the level of implementation of EBM in pilot MIHP sites. The recommendations on the obligatory implementation of EBM technologies are in MOH protocols but due to the absence of relevant indicators the current national reporting system does not let facilities to follow the effectiveness of implementation of evidence-based practices.

Therefore, thinking about the sustainability of effective MIHP monitoring system, there is a need to hand over the best practices of this system to MOH so as the essential data can be aggregated and used at different levels (maternity, oblasts, national). MIHP staff together with MOH statisticians, obg/gyns, and pediatricians has discussed and defined the new indicators to be incorporated into national reporting system. The format of some of the current MOH data collection tools (Delivery Log and Newborn Log) was changed to include these new indicators. However, the process of endorsement such changes might take time due to the national bureaucracy procedures. Therefore, the project has decided to probate the new scheme in several willing oblasts that are currently in the Project. The Project also will support the development of the database to be installed at oblasts statistic centers to facilitate the accumulation, analysis and making decisions on the data received from maternities and to transfer it to the national center of statistics.

<table>
<thead>
<tr>
<th>Themes 5: Capacity building for impact evaluation</th>
<th>Session number: [19.3]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session stream: Evaluating capacity development</td>
<td>Session name: Evaluating capacity development</td>
</tr>
</tbody>
</table>

**The Road to CAPACITY BUILDING for IMPACT EVALUATION: A SHOWCASE from the MENA Region**

*Presenter: Ziad Moussa*

As a Research and Development Center of Excellence in the MENA region, AUB-ESDU has always nurtured a keen interest in how evaluation can address complex development challenges.

ESDU has been actively engaged since 2005 with the Evaluation Unit at the International Development and Research Center (IDRC) in a number of projects aiming at building the capacities of ESDU to become a regional node on evaluation expertise while mainstreaming evaluation theory and practice in the MENA region.

One of the important projects executed by ESDU with funding from IDRC was the "Introducing and Exploring the Applicability of Outcome Mapping in the MENA region" project which held two conferences in Beirut (May 2007) and Cairo (April 2008) which brought together a very insightful group of M&E practitioners from various parts of MENA to discuss Outcome Mapping from the broader angle of M&E. Among the problems that were identified in the two workshop one could find:

- The absence of a regional M&E culture, which seems to be organically linked to the institutional governance problems that the region faces and which does not favor critical thinking.
- The weak accountability of the various stakeholders, which is reflected into a lack of proper monitoring and evaluation processes at the micro, meso and macro levels.
- The lack of a traceable M&E history at the level of the region, where the emphasis of projects and programs is on outputs rather than results and outcomes, and which translates into a weak system for generating indigenous M&E knowledge.
• Evaluation remains to a large extent a donor driven external exercise with weak ownership at the local level and the prevailing perception that M&E is mostly an external on-demand process.

In order to address these problems, ESDU has been putting together a set of activities centered on the “arabization” of Outcome Mapping and potentially other stakeholder oriented M&E approaches, initiating a community of practitioners on M&E in MENA, promoting the production, documentation and exchange of case-studies and best practices with a particular emphasis on expanding the knowledge base available in Arabic, etc…

<table>
<thead>
<tr>
<th>Themes 5: Capacity building for impact evaluation</th>
<th>Session number: [20.1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session stream: Evaluation networks</td>
<td>Session name: Evaluation networks</td>
</tr>
</tbody>
</table>

**Governance and Impact Evaluation: the Role of the Evaluation Networks in the promotion of a Culture of Evaluation:**

**Presenter: Pablo-Rodriguez-Billela**

The concept of governance has gained great popularity across most of the social sciences during the last decade, signalling a shift to a broad concern with a wide range of mechanisms with no presumption that these are anchored primarily in the sovereign state. Governance is broader than government, and it pays attention to the multiple ways in which governmental and non-governmental organisations interact, and to the ways in which political power and authority are distributed, both internal and external to the state.

Governance studies have lately stressed that assessing sustainable development interventions entails understanding not only what a particular project has been able to change, but also how the project itself was in the process impacted upon by the conditions under which it operated. Impact should be seen in the light of how beneficiaries appropriate encounters -by their participation- with external realities-, while at the same time generating meaning about the relationships which emerge out of it.

Studies of local governance are nowadays growing. They concentrate mainly on specific policy areas. Adopting a different perspective, in this paper I am interested in paying attention to the links between concrete practices and broader trends and processes in the relationship between the economy, the state, policy and society. Assuming a descriptive-analytical view of the governance perspective, I will argue about the usefulness of adopting an actor-oriented approach in order to shift the evaluation lens from looking at the outputs of local governance to looking at its processes -mainly from the perspective of those living in the local areas. In this way, a direct engagement with issues of participation will provide a key stimulus for a non-normative view of governance issues.

Specifically, I want to advance the discussion of local governance and the evaluation culture in the context of Latin America, where evaluation is many times –at best- considered a luxury, and other times a necessary evil. In order to exemplify the degree of institutionalization of the practice of evaluation, I briefly introduce the process of a recent sustainable evaluation of a national rural policy in Argentina which clearly shows the characteristics of a poor evaluation culture as well as the difficulties of turning sustainable evaluations into policy learning processes. Among other structural conditions, I will argue that the constitution of networks of evaluators is a positive way to deal with these challenges and to consolidate experiences of good governance. An illustration of this possibility is shown by introducing the Latin American Evaluation Network –ReLAC- and its role in the promotion and strengthening of a better culture of evaluation in the region.
What roles should professional networks play in the promotion of evaluation capacity building?

**Presenter: Jim Rugh**

The overall purpose of this “Perspectives on Impact Evaluation” conference is to enhance the demand for, quality and utilization of evaluations of the impact of development aid and development itself. Who are the persons expected to be adequately qualified and called upon to conduct such evaluations? Ideally, many if not most such evaluators should be professionals who practice within the countries where the evaluations are to be conducted (e.g., in Africa).

Those initiating evaluations have traditionally been international donor agencies. How do they identify qualified national evaluators? One possible answer to that question could be that a role that can be played by national associations of evaluation professionals would be to provide a clearinghouse where a list of adequately qualified evaluators is maintained.

Another important role of national evaluation associations is to promote and facilitate capacity building. There are a variety of ways to do so. One is to organize professional conferences where evaluators from the country as well as invited colleagues from elsewhere share their experiences and offer training in various frameworks and methods for conducting impact evaluations (as well as other forms of evaluation). Another role is to identify long-term and short-term educational and training opportunities within the country or elsewhere that are both appropriate to the needs of evaluators and accessible to them.

The main purpose of these proposed sessions in Cairo will be to solicit ideas and experiences from participants on what national associations of professional evaluators have and could do to promote such capacity building. Given the scope of the subject matter we are addressing, we are proposing two back-to-back (sequential) sessions:

1. A think tank session where any participating individuals will be invited to share their own personal experiences and perspectives—mainly how, from their perspectives, national networks have contributed to evaluation capacity building, or how they could do a better job of doing so;

2. A follow-up panel discussion where leaders of national, regional, and international evaluation associations/societies will be asked to share their experiences in organizing their groups to promote the demand for and practice of evaluation, and to respond to suggestions (and add their own) for how their groups could be more effective in promoting evaluation in general, and impact evaluation capacity development in particular. (We will try to identify such leaders prior to and during the Cairo conference.)

Some of the desired outputs from these sessions include:

1. A set of case studies (at least the initial ideas, to subsequently be followed up in more detail) for what national associations have done to promote capacity building;

2. Ideas for case studies on the role of evaluation associations in pooling local experience on actual impact evaluations undertaken;

3. Suggestions for how evaluation associations could play a role in promoting (encouraging, organizing) the demand for impact evaluations;
4. suggestions for what others (e.g. IOCE, AfrEA, NONIE, 3ie) could do to help strengthen national associations, enabling them to play more active and vital roles in such capacity building, especially for promoting and conducting impact evaluations; and

5. ideas for how national associations might benefit by partnering with other associations for purposes such as we'll be discussing during this session, e.g. other associations within Africa or beyond.

Themes 5: Capacity building for impact evaluation

<table>
<thead>
<tr>
<th>Session stream: Evaluation networks</th>
<th>Session number: [20.3 ]</th>
</tr>
</thead>
</table>

Evaluating networks: the challenge of a networked approach to evaluation

Presenter: Priyanthi Fernando

Managing networks, and participating in the evaluation of networks, has led to the recognition that the methodological challenges of capturing the effectiveness, efficiency, impact and sustainability of a knowledge network has very different challenges to that of evaluating a programme/project.

Knowledge networks typically engage in three types of activities:

- Information exchange and collaborative knowledge generation on the targeted issue (e.g. rural access or gender mainstreaming in energy policy)
- Moving knowledge into policy and action through improved communications and interaction with those who are in a position to put the research to use
- Managing the network in a way that network members’ capacity and that of the network to generate knowledge, communicate and facilitate change is strengthened. (Creech, Heather and Aly Ramji, 2004, Knowledge Networks Guidelines for Assessment Working paper, International Institute for Sustainable Development, Canada)

Knowledge networks have different structures, but commonly they strive to be democratic in their governance, and to ensure that members participate and learn from the networking activities. The process of networking is as important and as empowering as the outputs.

The International Forum for Rural Transport and Development (IFRTD) has, for instance, pioneered the ‘networked research’ approach to knowledge generation which “builds ownership, communication and advocacy into the overall design of the research programme” (IFRTD, 2006, A Networked Research Approach : a guide to conducting research in a network setting , IFRTD, London) and policy influence has long been a collaborative activity. Evaluations, especially external evaluations, unfortunately tend to take place within conventional programme/project evaluation frameworks and can miss the opportunity for the network itself to gain maximum benefit from the process or realize the true impact of donor funding.

There are many reasons why there isn’t as yet a networked approach to evaluation – there are methodological challenges as well as challenges in understanding the nature of networks and networking. Methodologies such as outcome mapping are taking us in the right direction but still require greater validation, especially from among donors.

In this roundtable, we aim to unpack and discuss some of these challenges and to begin the process for establishing “networked evaluation” as an accepted and useful methodology for both donors and networks.
M&E Networks and Associations; setting the agenda or following trends?

**Presenter: Mark Abrahams**

This paper reflects critically on the impact of networks such as NONIE, 3ie and others and more specifically on national and international M&E associations in relation to global and local demands for accountability and development success measures. The emergence of NONIE and 3ie is not merely a reaction to these demands but a concerted effort to influence, inform and to some extent dictate the M&E agenda at a global level as well as in local contexts. These initiatives are themselves informed by a critical need to provide evidence for the international multibillion development enterprise undertaken over the last few decades. However, the re-surfacing of impact evaluation, after its heyday during the 60s and 70s, follows hard on the heels of the Logical Frameworks, Outcomes Based Evaluation, Evidence Based Evaluation and more recently Results-based Evaluation. Each of these promoted vigorously by bi-and multi-lateral networks associated with the UN and the World Bank.

National associations, particularly young ones such as SAMEA (South African Monitoring and Evaluation Association), home to M&E practitioners from various disciplines, M&E users from different contexts and M&E commissioners with varying needs, are severely challenged in their quest to maintain their independence, act as platforms for good M&E practice, and satisfy their broad based members on very limited self generated budgets. These associations are often by-passed by international networks who by virtue of their bilateral relationships have access to government officials and other development structures in developing countries. And elsewhere, the AEA (American Evaluation Association) an established association has been compelled to react to their education department’s proclaimed preference for one approach to educational research, names RCTs, to assess impact.

From the vantage point of one national association, namely SAMEA, this paper argues that the institutionalisation of M&E and in this case impact evaluation depends largely on how people experience the State in the local context; how the State interprets and enacts its functions and the extent to which civil society in the broader sense is incorporated into this agenda. A purely methodological solution to engender accountability and ascertain attribution will also contend with the State’s multiple roles of governance, management and implementation. It is against this backdrop that the potential impact of the role of a national association such as SAMEA is explicated.

---

### The initiative of the Zambia Evaluation Association

**Presenter: John Ndlovu**

The proposal is on sharing the capacity building initiatives of the Zambia Evaluation Association (ZEA). It has played a major role in building capacity in impact evaluation through its participation on the Governance Monitoring Committee of the Government of Zambia (GRZ). Workshops on monitoring governance in Zambia for governance stakeholders have been conducted by Inwent German Centre for Economic, Environmental and Social Statistics. This was an initiative by ZEA and the Ministry of Justice. Stakeholders are trained in effective
monitoring and evaluating of the Firth National Development Plan. At the moment a governance survey is being planned.

The International Development Evaluation Association (IDEAS) had also been invited to give a lecture to M & E professionals on impact evaluation.

<table>
<thead>
<tr>
<th>Themes 4: Achieving influence: reporting, outreach, learning and policy impact</th>
<th>Session number: [21]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session stream: Evidence of policy influence</td>
<td>Session name: Evidence of policy influence</td>
</tr>
</tbody>
</table>

**Measurement on the Road to Impact: How do we assess the influence of evaluation and research on public policy?**

**Presenter: Fred Carden**

Evidence-based policy making is the mantra of many agencies. Research and evaluation are the sources of that evidence. However, affecting policy is not a simple matter of applying evidence and making a decision. It is a messy, uncertain business, fraught with contradictions, political and economic pressures and the need to influence over long periods of time before change happens. The road to influence can be a long one; and the road is winding and encounters many obstacles. This should not put us off efforts to influence policies, but we need ways to measure progress on the road to impact, not just the impact itself. Measuring efforts to influence policy is not the same as measuring the effects or impacts of an intervention once policy makers have determined to implement that change. When change comes from outside the policy process – when researchers and evaluators are trying to influence the agenda – measurement must start much earlier in the process if it is to help researchers and evaluators know whether, how and in what contexts they are having an influence. We have to look at measuring progress towards change, not only the change itself. How can evaluation support researchers to understand and affect policy influence processes? And how can evaluation help researchers strengthen their capacity to influence policy?

In a wide-ranging evaluation based on twenty-three cases, The International Development Research Centre (Canada) undertook to come to grips with how the research it supports influences policy agendas. Looking at cases in ICTs (information and communication technologies), in trade policy, financial reform, natural resources, social policy, mining and health and education reform, the study uncovered a number of important factors that researchers need to take into account to improve their opportunities for influence. The study has been recently published by Sage and IDRC: *Knowledge to Policy: Making the most of development research*.

The study identified stages of policy influence and also found that context is critical: what and how influence happens varies widely in differing contexts. There are strong parallels between the influence of research and the influence of evaluation on policy processes. Therefore, it is argued that these findings are important for the evaluation community beyond evaluating policy influence. The study identified five contexts and the factors that influence how policy is affected. Fred Carden will talk to the importance of these contexts in which researchers are trying to influence policy, and the implications for measurement. Drawing on the cases on information and communication technologies for development (in Uganda, Mozambique and South Africa), Zenda Ofir will speak to what this means in practice for increasing the use and influence of research and evaluation in policy processes.

The focus of this panel will be on the practical: what can evaluators and researchers take away that will help them re-focus their efforts to influence public policy? What are the parallels
between research and evaluation influence on public policy, and what are the distinctions? How can a multiple case method present evidence on the road to impact?

Themes 6: Impact evaluation of policies, programs, projects, partnerships and networks

Session number: [22.1 ]

Session stream: Governance and the rule of law

Session name: Governance and the rule of law


Presenter: Professor Victor A.O. Adetula

Democratic governance has recently become a major issue in development assistance policies, and also by extension an area of enthusiastic research in Africa. Western governments, donors, as well as international bodies such as the World Bank, are readily making their facilities available to government and non-governmental organizations towards ensuring good governance and democracy in Africa. This paper builds on author participation in the design and implementation of democracy and governance programmes in Africa as well as involvement in the conduct of evaluation and assessments in democracy and governance assistance projects. Essentially this paper critique and also seeks to improve on the dominant conceptual and methodological constructs on impact evaluations of assistance programmes in the democracy and governance sector.

There is today the sense of veritable crusade strategically organized around such themes like ‘democratic governance’ ‘good governance’, ‘democratic transition’ etc. Also, western scholars and researchers as well as development consultants working on the subject of democracy and governance in ‘new democracies’ and ‘transitional democracies’ are having a field day. A consequence of their intellectual efforts is the growing number of anthologies on the ‘failure’ of democracy in Africa. Using data from some assessments and evaluation reports on democracy and governance assistance programmes in Africa, and also some anecdotal evidences, we highlights the limitations and deficiencies in the dominant patterns in the designs, processes, outcomes and impact of bilateral and multilateral programmes of assistance to the democracy and governance sector in Africa’s new democracies. We take due cognizance of donors’ disappointment with low level of impact recorded in Africa’s new democracies and attempt to provide explanation beyond the assumptions and conclusions of assumptions and models based on experiences of western countries. In this paper we start by arguing that the transition that came with “wave of democratization” never ran their logical courses. Rather they were hurriedly replaced with foreign-aided reformist agendas with overbearing support of the international community. A consequence of these distractions is the emergence of fragile democracies that operationally can only display “a combination of both change and continuity”. Unfortunately many of the existing assessments and evaluation by donors and development partners have not fully appreciated the above and other structural contradictions as the key issues for consideration. Because of the weak governance environment, the expected ‘dividend of democracy’ has continued to elude the citizens despite routine elections and regime change in these countries. The masses in the new democracies are now responding with low level of loyalty, trust and support for the political system. We conclude that a new method of impact evaluation is required; one that is properly contextualized to address the limitations of the most of the existing general models and framework for democracy and governance programmes that are designs for western societies.
What do Anticorruption Programmes in Bulgaria Teach Us about Importance of Impact Evaluations?

**Presenter: Milena Minkova**

Donor-sponsored anti-corruption programmes have represented a large segment of international assistance to Bulgaria. Upon accession of Bulgaria to the EU on 1 January 2007, certain weaknesses that have remained in the areas of judicial reform, the fight against corruption and organised crime can prevent an effective application of EU-laws, policies and programmes and do not allow for Bulgarians to fully enjoy their rights as EU citizens. Therefore, further wave of anticorruption projects have been designed and implemented. The considerable international and national efforts invested in anticorruption activities in Bulgaria notwithstanding, corruption in Bulgaria continues to be characterized as “endemic” by the EU Monitoring Reports on the Cooperation and Verification Mechanism to assist Bulgaria to remedy these shortcomings and verify progress in the fight against corruption. Thus, while the ostensible objective of these programmes is to reduce corruption, progress in achieving objective stated has been limited, if any so far. The underlined discrepancy between efforts and results suggests that there are system problems with the very design, implementation and impact assessment of these projects that appear critical to their success. This paper addresses critical aspects of the above problems as a part of the broader effort to assess the effects of anticorruption projects in Bulgaria by looking at the specificity, relevance and fit of anti-corruption activities.

As a starting point, the paper posits that while projects seem to have generally succeeded in raising demand for reform, solutions to match that demand have yet to be found. The author questions both what reforms or change in particular the projects raised demand for, and what success the solutions applied thus far may claim. Emphasis has been put on the urgency to respond to both institutional and citizens’ needs. The author argues that the impact of reviewed projects was mostly of short duration, if at all. Projects generally failed to create a self-sustaining constituency to further their work, and when success was achieved it often depended heavily on contingent factors such as the presence of a “champion” or an exceptional level of donor resources targeted for a single, receptive client. The most successful projects provided direct benefits to a well-defined constituency. In all cases, the projects listed reducing corruption as one of their core objectives; yet based on interview material and project reports, no evidence has been found to support valid conclusions that projects had effectively reduced corruption.

In conclusion, the paper argues that project design has to be built on sustained public demand for a realistic, long-term anticorruption reform agenda. This should involve a shift from fighting corruption per se, the programme design that draws from a mix of broader discursive fields and the genuine attempt to improve project performance and decrease corruption - towards well targeted governance reforms that have a clear impact and benefits for those involved. Condition sine qua non for this move appear focused needs and impact assessment, without which the long awaited mobilizational potential of the anticorruption agenda might not be realized.
Challenges in Measuring Governance Impact: a case of ENCISS’ work in Sierra Leone

**Presenter:** Patrick E. Robin

**Co-authors:** Ahmed Ag Aboubacrine, Andrew Lavalie

Sierra Leone is a small country with very huge potential for development. It is endowed with vast fertile land for agriculture, huge deposits of minerals i.e. rutile, diamond, gold and bauxite. It has a population of approximately 5 million people. Despite its potential, 57% of its population live on less than one dollar a day and 26% live in extreme poverty. Over 70% of women and over 50% of the men are illiterate, average life expectancy is 42 years. A quarter of the children die before they reach their 5th birthday and mothers have a one in six chance of dying on childbirth. (This is considered the highest in the world).

Sierra Leone’s living standard steadily decline during the 1980’s. Much of the decline has been attributed to weak governance in the form of widespread corruption, dissolution of local councils, and ineffective rule of law. All this was exacerbated by the 10-year old conflict which destroyed properties and lives.

The acronym ENCISS stands for enhancing the interaction and interface between civil society and the state to improve poor people’s lives. ENCISS is a CARE-managed programme complementing DFID’s existing portfolio of support to Sierra Leone and GoSL, aiming to engage strategically with civil society and to contribute to an enabling environment for social and economic development.

This paper presents the challenges in setting governance impact indicators and in measuring them. It presents first what is measurable in governance (measurable aspects) and secondly how to measure that. Then an in-house experience of M&E for Governance is presented based on a framework that guides programmers to determine relevant governance impact indicators in various areas.

Governance is by nature a concept that applies to many sectors on which the impact is measured. Governance is not something directly viewable like health or education interventions. It takes time to see changes in governance, so what we measure is generally short-term outputs / results.

Governance impact is highly context specific, so harmonising governance impact indicators is a hard task. Thus, programmers have to use proxy indicators related to Process, Power, and Personal interests. There are some complex indicators and governance ME systems (Transparency International Index, Afrobarometer, Mo Ibrahim Index, NEPAD/APRM) which are not always useful in measuring the impact of local interventions.
The Unrealized Potential of Survey Research as an Impact Measurement Tool in Democratic Development and Governance

Presenter: Margaret Sarles

Co-authors: Devra Moeller, David Black, Mark Billera

USAID and other donors have provided extensive support to survey groups such as Afrobarometer and AmericasBarometer, to carry out democracy and governance surveys across nearly all of Latin America and major countries in Africa. Some countries have been surveyed at least six times at regular intervals, using many of the same questions over the entire period. USAID missions commission or contribute to funding them for a number of reasons. They capture underlying characteristics of the political culture; they ascertain political activities and beliefs; they measure levels of individual-level corruption and other broader political characteristics of the country; they are a major tool of dissemination and reform, as results are highly publicized and considered important by political stakeholders. Within USAID missions, they are also a major source of evaluation data, with the rigor and periodicity needed to measure change accurately.

Over the last five years, particularly with AmericasBarometer in Latin America (and to a smaller degree in Africa), we have introduced over-sampling into the mix to begin to develop a data base that allows impact evaluation through testing the counter-factual. A number of missions are now doing this, particularly in the area of decentralization and participation, although in other aspects of the democracy and governance portfolio as well.

The paper will lay down the theoretical case, and the policy argument, on why and how democracy surveys are and should be an important tool in impact evaluation. It will give examples of specific USAID missions in Latin America and the Caribbean that have extensively used survey research results in evaluation and are making or attempting to make use of them in developing impact evaluation tools.

It will then assess the progress that has been made, and the obstacles that are also increasingly clear for maintaining a focus on over-sampling and impact analysis generally. Some missions are expanding their use of surveys as an impact evaluation tools, others are abandoning it. The paper will review specific experiences from the field. It will also include the perceptions of the survey groups in Latin America and Africa who have worked with missions on introducing, or trying to introduce, impact methodologies. We will analyze the issues surrounding impact evaluation through survey data both at the country level and within USAID itself – the bureaucratic incentives and disincentives for long-term, rigorous evaluations.

Finally, the paper will suggest specific technical and policy recommendations abouth the field/country level and within the broader survey and donor communities that would let us better "realize the potential" of the surveys as as an impact evaluation tool.
The impact of social investment funds on social capital and violence

**Presenters: Simon Brook, Luca Pellerano**

A prominent auxiliary goal of development and poverty reduction is to strengthen community networks and improve state-society relations in developing countries through the promotion of participatory institutions and processes. One of the primary vehicles of these efforts over the last two decades has been the Social Investment Fund (SIFs). The basic idea is that supporting and strengthening local capacity for design, implementation and maintenance of social and economic infrastructure will also increase ownership of development projects, promoting their sustainability and effectiveness. SIFs are small, flexible, special programmes organized to make social investments in poor communities. Typically, they are "demand-driven," based upon a predefined menu of projects with well-established eligibility criteria, whose actual application and specific investment mix is predetermined by the presentation of requests from eligible groups. Departing from traditional approaches to development, they enable communities and local institutions, rather than central governments, to take the lead in identifying and carrying out community-level investments.

Social Investment Funds have been object of several impact evaluation studies in recent years. The majority of them have focused on understanding the effects of SIFs on traditional measures of well-being and multidimensional poverty. This paper systematizes the existing information on the effect of FTIs on social capital and violence, with the objective of identifying lessons learned that can be drawn based on the existing evidence and determining an agenda for further evaluation in the future.

**Fostering Social Cohesion? A Synthetic Review of Social Cohesion Interventions in Africa**

**Presenters: Elisabeth King, Cyrus Samii**

This paper examines development and reconstruction interventions in sub-Saharan Africa aiming to generate social cohesion. The goals of social cohesion programs include promoting institutions and relations necessary to manage inter-individual and inter-group conflict—the social capital dimension—as well as attitudes reflecting a willingness to move beyond zero-sum or vengeful relations—the reconciliation dimension. A pro-social cohesion orientation has informed major development interventions in Africa. The review synthesizes the findings from evaluations of recent social cohesion interventions. Our aim is to contribute to evidence-based policy-making and to a scholarly research agenda. The key questions our review tackles are as follows: What development projects in Africa have been studied in terms of their impact on social cohesion? What did the studies find? In particular, what are the mechanisms underlying the production and destruction of social cohesion? What can we learn about scope conditions? What can we learn about appropriate measures?
The Effect of Education on in Prison Conflict: Evidence for Argentina

**Presenter: Maria Laura Alzua**

Even though prisons have existed since ancient times, before the late eighteenth century, they were not commonly used by justice systems around the world for punishing criminal behavior. In modern days the situation is quite the opposite and most serious felonies are often punished with prison.

Two of the most cited justifications for such punishment in modern society are reductivism and retributivism. The former is a forward looking theory associated with the idea that through deterrence, incapacitation and reform of prisoners, future violent and criminal behavior can be reduced. The latter on the contrary, is a backward looking theory based on the idea that criminals should be punished because they deserve it, relating such ideas to the theory of "social contract" in which punishment is a mean through which equality is restored.

The economic literature has largely studied and justified punishment under the reductivism channel, especially through its effect on deterrence. Evidence shows that through an appropriate choice of policy variables such as expenditure on police and courts, the expected returns of crime to potential felons would decrease reducing its incidence.

However, much less focus has been given to the effect of the penal system on the "reformation" of prisoners. Among the few programs established in modern societies to try to reform convicts are education programs. On this matter, the economics of crime literature has acknowledged several channels through which education may affect criminal participation. In particular, once the prisoner is released, education by increasing potential wage rates can reduce the probability of engaging in criminal activities. Moreover, schooling may affect directly the financial or psychic reward of crime since it can affect behavioral attitudes, time preferences or risk aversion indirectly influencing the decision to engage in crime. Finally, education could affect crime through the fact that past conviction can hurt subsequent labor opportunities of the more educated due to stigma.

Using census data for Argentine prisons this paper is a first attempt to estimate the direct impact that prisoners education programs have on violent behavior within prisons. Specifically, we evaluate whether participation in basic education programs reduces the probability of in-prison conflicts. To handle the possible endogeneity of schooling decisions an instrumental variables approach is explored. We find that mandatory primary and secondary education participation lowers the probability of in-prison conflict of about 20 percentage points on average for different measures of in-prison conflictivity using probit and linear probability model specifications.

This study represents a contribution supporting the idea of reductivism, where prison's objective should be to reform inmates. Furthermore, evidence provided suggests that fostering in-prison education could have a positive effect in reducing in-prison conflictivity. This result could be important for policy actions in different fronts. Most importantly, education programs could improve the quality of life of prisoners and increase their human capital. This in turn could bring two positive implications. On the first hand evidence in the economics literature has found that education provides ex-convicts with better working opportunities in the future. Secondly, recividism could also decrease given that prisoners with lower misconduct have lower subsequent criminal activity.
Health Programme Evaluation by Propensity Score Matching: Accounting for Treatment Intensity and Health Externalities with an Application to Brazil

Presenter: Rodrigo Moreno-Serra

Most of the literature on health programme evaluation has estimated average programme impacts relying on either: (i) data on the presence or absence of an intervention in a particular locality, or (ii) data on individual participation in the health programme. By estimating an average health impact which is independent of the programme's population coverage, the empirical approaches of these studies overlook the important fact that public health interventions create externalities whose magnitude depends crucially on the number of covered individuals in a locality. Recent research has convincingly shown that treatment externalities derived from health programmes exist and can be large enough to severely bias programme impact estimates if not taken into account. However, in general, most existing impact evaluations are not well-equipped to incorporate these health externalities into their estimated average treatment effects, and their separate measurement is clearly difficult without treatment randomisation at an adequate level. The main contributions of this paper are to suggest and apply an empirical approach for the impact evaluation of public health interventions which also takes into account treatment externalities--therefore providing a comprehensive account of the total health benefits due to a given intervention--when only non-experimental, routine data are available. The proposed framework involves the computation of average treatment effects by a propensity score matching-difference-in-differences estimator adapted to the case of multiple treatments, jointly evaluating the impact of different programme coverage levels. Any programme health-related externalities accruing to individuals who do not directly receive the programme services will be included in the estimated average treatment effects by using these coverage levels as the treatment variable of interest. The methods are used to conduct an impact evaluation of the Family Health Programme (Programa Saude da Familia), the broadest health programme ever launched in Brazil, on adult and child health, with data from the largest national household survey.

Housing Health and Happiness

Presenter: Sebastian Martinez

Co-authors: Cattaneo, Galiani, Gertler and Titiunik

Despite the importance of housing for people's well-being, there is little evidence on the causal impact of housing and housing improvement programs
on health and welfare. In this paper, we help to fill this gap by investigating the impact of a large-scale effort by the Mexican Government to replace dirt floors with cement floors on child health and adult happiness. We find that replacing dirt floors with cement floors significantly improves the health of young children. Specifically, we find significant decreases in the incidence of parasitic infestations, diarrhea, and the prevalence of anemia, and an improvement in children’s cognitive development. Additionally, we find that replacing dirt floors by cement floors significantly improves adult welfare, as measured by increased satisfaction with their housing and quality of life, as well as by lower scores on depression and perceived stress scales.

Themes 2: Impact evaluation: Approaches and methods
Themes 6: Impact evaluation of policies, programs, projects, partnerships and networks

Session stream: Health

The Value of Measurements and Where they Apply: Monitoring Requirements for Evaluation of Population Policy

Presenters: Liezl Coetzee, Ina Mentz

This presentation will examine some of the challenges experienced in evaluating the impact of South Africa’s Population Policy which was adopted in 1998, in alignment with the principles of the International Conference on Population and Development, which took place in Cairo, Egypt, in September 1994. The policy shows a shift in emphasis from the former focus on issues around fertility, mortality and family planning, with quantitative demographic targets used to measure success, to a new focus looking at the impacts of demographics on development, using qualitative indicators related to the improvement of standards of living. The shift in emphasis presents particular challenges when it comes to monitoring and evaluating impact of the policy, which essentially requires measuring or assessing human development.

The presentation will review a series of policy assessments that have been undertaken since 1998, looking at approaches used to measure impact. A most recent evaluation is currently underway, and challenges encountered in this process will be highlighted.

One of the most significant challenges noted relates to the availability, as well as reliability and comparability of data and information on population and development over time. In this regard the important link between ongoing monitoring and periodic evaluation will be highlighted with reference to the policy objective to “ensure the availability of reliable, up-to-date information on population and human development situation to inform policy making & programme design, implementation, monitoring and evaluation”

Other challenges that will be noted include:

- Management of policy implementation relating to the various roles and responsibilities of multiple governmental and non-governmental stakeholders in implementing the policy
- Attributing changes to the policy. While changes may be noted in “population and human development trends” in South Africa since the 1990s, the Population Policy cannot be claimed to have caused such changes. Indicators such as the Human Development Index are very useful to see changes in trends, but they cannot be used to determine the causes of change. This is further complicated by the fact that different Departments, as well as
numerous NGOs, collaborate to address a range of social ‘problems’, thus making it difficult to isolate the contribution of a specific department to the solution of a problem (or not)

- Defining ‘success’, noting the example of measures to address the issue of HIV/AIDS. Despite efforts of different departments (& other role players) to curb the problem, the problem keeps growing. This raises questions about the effectiveness of various programmes and initiatives, though one can also ask how much worse the situation might have been without these initiatives.

Themes 2: Impact evaluation: Approaches and methods
Themes 6: Impact evaluation of policies, programs, projects, partnerships and networks
Session number: [23.1]
Session stream: Health
Session name: Health

Public Health Insurance for the Poor: Targeting and Impact of Indonesia’s Askeskin Programme

Presenter: Robert Sparrow
Co-authors: Asep Suryahadi and Wenefrida Widyanti

Health care utilisation and public spending in Indonesia falls behind its South-East Asian neighbours, while inequality in health care utilisation is relatively high. Out-of-pocket (OOP) health payments are also relatively low compared to other Asian countries, but this apparent low propensity to spend is accompanied by a high variation across the population, with the non-poor allocating a larger share of their budget on OOP spending.

Health shocks and the subsequent cost of required health care can have implications for both transient poverty and long term poverty traps if households are resource and credit constrained. In the absence of access to credit markets, households may employ alternative coping mechanisms. But while idiosyncratic shocks seem insurable, full insurance is often not feasible. Moreover, when faced with covariate shocks and chronic illness, coping mechanisms are ineffective and informal insurance fails.

Within this context, we investigate Indonesia’s Askeskin health insurance programme for the poor, which was initiated in 2005. Under this programme, public health insurance was targeted to the poor, entitling the recipient to free basic health care at public health clinics and inpatient services at public third class hospital wards. In addition, Askeskin coverage includes an obstetric service package, mobile health services and special services for remote areas, immunisation programmes, and medicines. With a budget of USD 400 million for 2005 Askeskin aimed to cover 60 million people.

In this paper we will investigate the implications of Askeskin for access to health care for the poor. We will first analyse targeting performance of Askeskin in relation to socioeconomic status of recipients, actual OOP payments, and exposure to OOP payments given households’ health care needs. We then proceed with estimating the impact of Askeskin on outpatient and inpatient utilisation.

The analysis is based on a longitudinal socioeconomic survey (Susenas panel) conducted in 2005 and 2006 for a nationally representative panel of 8,582 households. The first wave of this survey was conducted just before start of Askeskin, providing a baseline. Identification of treatment effects therefore relies on a difference-in-difference approach.
Askeskin has been successful in targeting the poor, despite some non-trivial leakage to the non-poor. In addition, Askeskin seems to have been allocated proportionally more to individuals that live in households that are expected to require a relatively high OOP health care budget share in order to meet health care needs.

There appears to be a strong impact of Askeskin, as it increases both inpatient and outpatient care. The impact is particularly strong among the non-poor, suggesting that there remain barriers to health care utilisation for the poor that are not overcome by Askeskin. Hence, lack of health insurance seems only partly responsible for under-utilisation of health care among the poor. Other explanations could lie with indirect and opportunity costs of seeking health care.

Evaluating Community Participation for Maternal Health in Developing Countries

**Presenter: Lucia D’ambruoso**

Community participation in health is an under-researched, yet key, area of public health and development. Participatory principles state that the insights of those directly involved with a good or service should be central to the processes of design, implementation and evaluation. Community participation was enshrined as a fundamental principle in the Alma Ata Declaration of primary healthcare in 1978. Community participation can be defined as a capacity-building process through which community individuals, groups, or organizations plan, carry out, and evaluate activities on a participatory and sustained basis to improve their health and other needs, either on their own initiative or stimulated by others.

Despite the political rhetoric and some promising ‘evidence’ of effect, a distinct lack of clarity on concepts, processes and terminology persist. A poor theoretical basis has also led to a paucity of evaluation methods. Community participation does not conform to the input-process-outcome format of health impact evaluation. It is process-based, dynamic, and highly context-dependent. As a result, some scholars assert that generic evaluative methodologies are unattainable because participation is the result of specific conditions and situations and so participation, and evaluation of its health effects, in the absence of grounded experience is not possible.

This paper presents the findings of a synthetic review of community participation for maternal health in developing countries. Two elements are presented. Firstly, a review of approaches to conceptualise community participation is provided. This includes a consideration of the biomedical, sociological and inter-disciplinary theoretical perspectives and underpinnings. Secondly, the evidence of effect of community participation for maternal health in developing countries and the mechanisms by which effects are detected are presented. The paper considers whether the benefits of community participation are amenable to normative healthcare evaluation, and if not, what the alternatives are.

In the absence of a unifying framework, careful and systematic documenting of the processes of community participation is urgently needed to identify and share best practices. This paper explores how, and whether, community participation for maternal health in developing countries works and explores the potential for establishing a firmer foundation for the implementation and evaluation of this promising approach for maternal health in developing countries.
Neonatal period is recognized as a brief, critical time that requires focused interventions to reach the MDG (two-thirds reduction in child mortality) by 2015. In India, there are one million neonatal deaths every year, representing approximately a quarter of all global neonatal deaths.

To cater the need, Integrated Nutrition and Health Programme (INHP) was launched in 100 districts in India with partnership of an international NGO, CARE-India, together with Government of India (GOI) and local NGOs. The programme was implemented through the infrastructure of GOI Ministry of Women and Child Development’s Integrated Child Development Services (ICDS) and the Ministry of Health and Family Welfare. The newborn care package aimed to increase the frequency of behaviours during the antenatal, delivery and postnatal periods that have proven benefits for maternal and newborn survival.

**Objective:** To assess the impact of the newborn health component of a large-scale community-based integrated nutrition and health programme.

**Methods:** Using a quasi-experimental pre-post design, we evaluated the INHP by selecting two rural districts of Uttar Pradesh, India. One INHP district, Barabanki, served as the intervention district, while a district receiving standard government health and ICDS, Unnao, was used as a comparison district. From the 15 rural blocks in each district, we used a computer programme to randomly select nine blocks in the intervention district and eight blocks in the comparison district. One sector, an area with 15–25 anganwadi centres and an estimated population of 20 000 to 25 000 people, was randomly selected from each of the selected blocks. The sample size was calculated to detect a 20% reduction in neonatal mortality following the intervention with 80% power at 5% significance level.

Mothers who had given birth in two years preceding the surveys were interviewed during the baseline (n=14,952) and endline (n =13,826) surveys in 2003 and 2006, respectively. The primary outcome measure was reduction of neonatal mortality.

**Findings:** In the intervention district, the frequency of home visits by community based workers increased during the antenatal (from 16% to 56%) and postnatal (from 3% to 39%) periods, as did frequency of maternal and newborn care practices. In the comparison district, no improvement in home visits was observed and the only notable behaviour change was that women had saved money for emergency medical treatment. Neonatal mortality rates remained unchanged in both districts when only an antenatal visit was received. However, neonates who received a postnatal home visit within 28 days of birth had 34% lower neonatal mortality (35.7 deaths per 1000 live births) than those who received no postnatal visit (53.8 deaths per 1000 live births), after adjusting for sociodemographic variables. Three-quarters of the mortality reduction was seen in those who were visited within the first 3 days after birth. The effect on morality remained statistically significant when excluding babies who died on the day of birth.
**Conclusion**: The limited programme coverage did not enable an effect on neonatal mortality to be observed at the population level. A reduction in neonatal mortality rates in those receiving postnatal home visits shows potential for the programme to have an effect on neonatal deaths.

**Themes 2: Impact evaluation: Approaches and methods**

**Themes 6: Impact evaluation of policies, programs, projects, partnerships and networks**

**Session number**: [23.2 ]

**Session stream**: Health

**Session name**: Health

**Comment les facteurs contextuels relevés par l’analyse de l’implantation peuvent-ils aider l’évaluation d’impact subséquente: Exemple de la stratégie de prise en Charge Intégrée des Maladies de l’Enfance (PCIME)**

**Presenters**: Leon Nshmyumuklza, Marie Gervais

La considération des facteurs contextuels est un élément important pour assurer la validité d’une évaluation, plus particulièrement celle d’impact qui vise à démontrer les effets attribuables à l’intervention. Ces facteurs peuvent être de nature socio-économique, démographique, environnementale ou autre. En effet, les facteurs contextuels peuvent être à la fois confondants et modifiants. Confondants dans la mesure où ils peuvent créer la confusion entre l’intervention et son impact potentiel sur la santé. Modifiants dans le cas où ils peuvent modifier les effets de l’intervention et ainsi affecter la généralisation des résultats. Ceci peut être particulièrement important pour des programmes de santé pour lesquels l’impact peut varier de façon substantielle suivant le contexte d’implantation. La compréhension des éléments contextuels obtenus lors de l’analyse de l’implantation de l’intervention s’avère alors important non seulement pour améliorer la validité de l'évaluation d'impact ultérieur mais encore peut aider à prédire l’impact que pourrait avoir cette intervention.

Cette présentation vise à montrer, à partir de l'exemple de la PCIME, comment la connaissance et l'analyse des facteurs liés à l'implantation est très importante pour l'interprétation des résultats de l'évaluation d'impact. La PCIME une stratégie conjointe de l’OMS/UNICEF dont l’impact est l’amélioration de la santé et de la nutrition ainsi que la réduction de la mortalité des enfants de moins de cinq ans dans les pays à ressources limitées. Nous allons présenter les facteurs contextuels que nous avons relevés lors de notre évaluation d’implantation de cette stratégie dans un district du Rwanda ainsi que la stratégie méthodologique utilisée pour les documenter. Nous allons également examiner comment ces derniers peuvent venir influencer l’évaluation d’impact subséquente.

À la fin de cette présentation, les participants seront capables de comprendre comment les facteurs contextuels doivent faire l’objet d’une attention particulière et qu’il est nécessaire de prévoir comment ils seront objectivés lors d’une évaluation. Cela constitue une étape incontournable si on veut s’ass
Evaluating the effectiveness of behavior change interventions

**Presenter:** Nadra Franklin

**Co-authors:** Kellie Kim, Elisabeth Sommerfelt, Zo Rambelson, Mellisa Panagides-Busch, Audrey-Marie Schuh Moore

Using case studies of multi-country approaches and multi-method evaluations, this group of evaluators will present 4 examples of evaluations of behavior change interventions in public health and education. We will 1) describe methods for designing and implementing a comprehensive evaluation for educational reform in Egypt, Ethiopia, Liberia, and Zambia; 2) discuss using large-scale household surveys, cost and effectiveness analysis, and DALY calculations to measure the impact of a global infant feeding project; 3) present examples of using propensity score designs in health and education interventions; and 4) provide examples of using existing, secondary data for impact evaluations.

Taken as a group, these 4 case studies present a diversity of methodological approaches to measuring impact in multiple global settings. We will describe a comprehensive evaluation approach to measure educational reform that included both qualitative and quantitative measures. We will describe the rationale for implementing annual, large household surveys during a long-term infant-feeding intervention, and discuss how those data were used to estimate cost effectiveness based on specific behavior change indicators. Compared to this intensive approach, we will review the use of propensity score analysis at the end of projects and review the use of this method for measuring impact. We will also consider the use of secondary data to measure program impact, on its own and paired with primary project data.

Examples come from country activities in Bolivia, Ghana, Egypt, Ethiopia, Honduras, Liberia, Madagascar, and USA.

Combined Longitudinal Qualitative and Quantitative Monitoring for Results as part of an Integrated Evaluation Strategy

**Presenters:** Jennifer Robinson

The presentation will examine the process and principles applied in the development of a longitudinal monitoring and evaluation system that meets both donor reporting and internal management information requirements through the delivery of high quality data that aligns to a result based management strategy. Data quality parameters will be discussed as an integrated part of the presentation. This presentation will highlight the importance and
advantages of longitudinal monitoring and evaluation of results as part of an integrated evaluation strategy.

We will start with an overview of the results based management framework and the alignment of the framework to the theory of change underpinning the implementation. The presentation will progress through the participative process of developing qualitative indicators that measure results/impact at the short, medium and long term. We will take a brief look at best practice in the measurement of indicators and management of data at each step in the data management cycle.

We will examine the use of an adaptation of the “Most Significant Change” method of collecting qualitative data on a longitudinal basis, and using this data to identify and measure both anticipated and unanticipated change at different levels in implementations.

The importance of aligning data management processes to normal daily activities of programme staff, particularly at source and collection, and the resultant impact on the quality of data will be highlighted. Perhaps the most significant challenge to the implementation of longitudinal monitoring and evaluation systems is the anticipated additional administrative burden these systems imply. We will briefly look at methods of minimising administration associated with longitudinal data management processes.

The presentation will highlight significant challenges encountered during development and implementation as well as best practice and results highlighted as a result of system implementation.

Principles discussed will be illustrated using the monitoring and evaluation system implemented by the Harvard School of Public Health’s (HSPH) Centre for the Support of Peer Education (CSPE) Programme in South Africa and adapted for implementation by the Anglican AIDS and Healthcare Trust’s Siyafundisa Programme.

---

**Effectiveness of Community Participation in Tuberculosis Control Program at the Eastern Cape Province, South Africa**

**Presenter: Tom Achoki**

**Introduction:** With the global prominence of Tuberculosis (TB) as a public health issue, various multi-stakeholder interventions have been adopted to meet the challenge. In low resource settings where health systems are hardly coping with the disease burden, community participation has emerged as a pivotal measure for successful programming. For this study the community intervention package comprised of patient and community sensitization employing multimedia messages, community gardens and facilitation of treatment supporters among other community based initiatives. The study objective was to determine the best approach of integrating community interventions in tuberculosis programming in low resource settings.

**Methods:** This evaluative study considered all the 3110 new TB patients registered in three sub districts in Chris Hani District, Eastern Cape Province, from quarter 1 2004 to quarter 4 2005. The performance of the three Tuberculosis programs was compared over time; taking cognizance of the community intervention in one of the sub districts (Sakhisizwe) during the said period. Further analysis was done to establish the influential determinants of treatment success.

---
**Results:** Bacteriological coverage in the interventional sub district dropped, while the controls remained consistent at around 90%. Similarly the new smear conversion rates in the interventional sub district declined from 80% to 50%. The new case cure rates followed suit, with the interventional sub district reporting a declining from 60% to 30%; while one of the controls improved to register 80% cure rate. Both the intervention and control sub districts registered a drop of defaulter rates to around 10% over the same period. However, the proportion of patients reported as “not evaluated” increased in the interventional sub district while the controls registered a drop. Patients registered in the clinics had better chance of successful treatment outcome (OR 10.8, 95% CI 8.03-14.3) compared to their hospital counterparts. Similarly newly registered Tuberculosis cases performed better (OR 1.48, 95% CI 1.2-1.75) in than their transferred counterparts.

**Conclusion:** Community participation by itself is not adequate to improve the performance of a poorly performing TB Program. It is important to enhance the technical and organizational capacity of the program, such as quality diagnosis, case holding and patient evaluation as well as recording and reporting; prior to engaging community approaches. Failure to observe this logical relationship would ultimately result in disastrous consequences, since a TB Program with technical shortfalls is not likely to adequately meet the increase in service demand. Therefore the process of entrusting communities with more responsibility in a TB Program should be gradual and take cognizance of the various contextual and health system factors.

---

**Themes 2: Impact evaluation: Approaches and methods**

**Themes 6: Impact evaluation of policies, programs, projects, partnerships and networks**

**Session number: [23.4]**

**Session name: Health**

---

**Independent Evaluation of National Leprosy Eradication Program (NLEP) in India**

**Presenter: Sutapa Agrawal**

India is a home of highest number leprosy patients (92005 cases in 2007) in the world. Government of India (GOI) has initiated NLEP by integrating leprosy with general health services. The Independent Evaluation of NLEP in India was carried out by IIHMR with financial assistance from WHO during February–May 2008. The study attempted to evaluate the quantitative and qualitative achievement made under different components of NLEP and gain a complete understanding of the programme impacts with the following objectives:

- Progress of NLEP in terms of programme coverage
- Achievement based on selected indicators
- Integration of leprosy services
- Management of drug supplies
- Quality of data and reporting
- Role played by NGO

The inputs provided by the study are expected to form a solid database and direction for future planning and initiatives related to eradication of leprosy in India.
The African Youth Alliance (AYA), a partnership between PATH, Pathfinder International, and the United Nations Population Fund (UNFPA), worked in four African countries from 2000-05 to improve adolescent sexual and reproductive health (ASRH) through integrated, comprehensive, and multi-sectoral programs. The project, which was funded by the Bill and Melinda Gates Foundation, employed a multi-faceted approach that included policy and advocacy for a supportive ASRH environment; effective behavior change communication for ASRH; integrating ASRH into existing livelihood skills development programs for youth (in-school and out-of-school); and expanding access to youth friendly services. The approach was unique in seeking to implement all components simultaneously, while building capacity and fostering coordination among established partners to encourage sustainability of ASRH services.

In 2005, as AYA entered its final year, the Gates Foundation asked John Snow, Inc. (JSI) to carry out an impact evaluation of AYA to determine whether the project had produced measurable impact on ASRH behavioral outcomes among young people in AYA countries, and whether the change in outcomes could be attributed to AYA interventions. To carry out the evaluation, a population-based survey was carried out in 2006 among youth aged 17-22 in Ghana, Tanzania, and Uganda to evaluate exposure to and impact of AYA in those countries. The post-test only research design determined program impact by comparing behavioral outcomes 1) between intervention sites and controls, and 2) between youth who were exposed to AYA programs and those who were not. A combination of propensity score matching and instrumental variable approaches was used to analyze results. The main outcome indicators included age at first sex, practice of abstinence, number of sexual partners in past 12 months, ever use and current use of contraceptives, ever use of condoms, consistent use of condoms, and condom use at last sex. Antecedents to behavior change such as knowledge, attitudes, and self-efficacy were also measured and analyzed. Two-stage cluster sampling was employed to carry out the household surveys, with sample sizes of 3,416 in Ghana, 1,900 in Tanzania, and 3176 in Uganda.

The evaluation determined that AYA had statistically significant positive impact on ASRH outcomes among female youth in all three countries, with the strongest impact related to condom and contraceptive use. Impact among females was weaker for abstinence and sexual debut. Impact was also notably weaker among males; in fact, in Ghana and Uganda, there was no significant positive impact on any behavioral outcomes among males. In this presentation, we will discuss the main elements of the AYA program, the objectives and design of the impact evaluation, the analysis methodologies, key findings in terms of outcome and antecedent indicators, discussion of implications, and recommendations. It is hoped that findings from this evaluation will provide valuable insights for improving future ASRH programs in AYA countries and other African settings, both in terms of programmatic design and evaluation methodologies.
Thinking spatially about Health Inequalities: Developing a Spatial Evaluation Framework to Address Social and Health Problems

Presenters: Derek Powell, Sanjeev Sridharan and Julia Koschinsky

Background

This think tank aims to build understanding of how evaluations informed by a spatial framework can contribute to addressing problems of health inequalities (and improving things for those with greatest need).

This think tank is a dialogue between policy leaders in Africa with evaluation centres in Canada and U.S. Building on existing research partnerships, this project can potentially lead to insights arising from innovative spatial based evaluation frameworks to develop better solutions to address problems of health inequalities in developing countries.

Part of what motivates our collaboration interest is a search for better ways to create evaluation systems that improve the linkage between a systematic analysis of health problems to solutions that incorporate place-based thinking. This motivation also drives a larger current research effort and movement towards "evidence-based policy," "turning statistics into knowledge," and "measuring progress in societies." Our approach would be to explore how far the applied evaluation approaches, methods, and tools turn out to be useful - rather than assuming a priori that they will be useful.

Key to our thinking has been recognition of the differences between “problem space” (knowing what variables or systems of relationships are associated with the problem) and "solution space" (the strategic drivers with scope to reverse the problem). We are interested in exploring the role that spatial thinking can play in building bridges between the problem and solution space. We are also interested in exploring if innovative spatial tools can provide insights on reducing health inequalities by taking into account the heterogeneous needs of individuals and communities.

The focus of the presentation will be on the following:

- how to use spatial tools to help understand and tackle development problems in a context of massive spatial inequality;
- the ways in which the analysis can help us to decide the mix of policy interventions that will make a lasting impact on people's lives given high rates of poverty and finite resources.

Examples of policies that we will discuss include: The National Spatial Development Perspective; Redistributive social expenditure; Integrated Rural and Urban Development Strategies; and Municipal integrated development plans.

Presenters: Charlie Teller, Assefa Hailemariam and Yared Mekonnen

This paper will examine the experience, progress and challenges in generating the demand for, production of and use of rigorous impact evaluations in the government’s high priority areas of population, food security, health and development policies. The authors share their vast experience in government, university, think tank and international development organizations in evaluation research and rigorous policy, program and project evaluation over the past 15 years.

This has been a slow process since the fall of the Military Regime in 1991 and the extreme poverty and food insecurity in the 1980s and 1990 in developing a societal and organizational culture of reliable and trusted information to inform data-challenged policy makers. The incipient demand toward more objective, accountable and transparent production and use of evidence has been recently aided by government-international partnership commitment to Millennium Development Goals, Strategic Planning for high profile development policies, as well as to capacity building in monitoring and evaluation at major universities around the country.

Three high level policies and strategies will be examined in terms of their impact evaluation:

1- The National Population Policy of 1993 and subsequent Plans of Action

2- The Poverty Reduction and Food Security Policies and Strategies (3rd one started in 2006)

3- The Health Sector Policy and Sector Development Program, 1998-2011

A recent assessment by the authors of the demand, access and use of population, health and development data for decision-making will inform this paper. Also, the process of capacity building in the development of strategic information, sectoral information management and M&E systems, as well as training for evaluation, the creation of an African Center of Excellence in Addis Ababa, and the proliferation of local M&E expertise will also be documented. Finally, constraints by policymakers in their demand for, access to and use of impact evaluations will be discussed, with recommendations for bridging the gap between evaluation research and policy, and the proper use of more objective, timely and transparent impact findings.
New ideas for implementing impact evaluations of behavior change interventions and projects

*Presenter: Jennifer Barker*

*Co-authors: Nadra Franklin, Kellie Kim, Zo Rambeloson, Holly Ladd, Joshua Volle, Veronica Lee, Martin Alilio, Alexandra Fallon*

We propose a round table discussion on new ideas being implemented on the ground to make impact evaluations more inclusive of community partners, supporting the generation of high quality data over a shorter timeframe, working with local evaluators, and using GIS mapping for both evaluation data and to help with intervention targeting. Topics that will be included (but not limited to) are:

- Working with local community partners to measure outcomes at the project and program level
- Identifying and measuring impact on beneficiaries
- Identifying and measuring impact on systems
- Using ICT and GIS mapping tools for faster, higher quality, and focused data collection

In an HIV and AIDS project working with CBOs in Namibia, we have trained CBO members to gather behavioral data that can guide them and give them an idea or suggestion of the direction in which their community is moving towards the identified target behaviors and norms. Not only does this approach capture a different “type” of impact data, but it has an unintended impact on the CBO members collecting and interpreting the data. We will discuss our attempts to measure program impact on the systems in which our programs are implemented using network analysis and systems thinking. As a solution to the delay in accessing data and the need to translate numbers into visual maps, we will discuss the use of ICT for routine and survey data collection in multiple countries with multiple partners, and the use of mapping, especially of infectious disease.

*Health Insurance in Nicaragua: Enrollment, Retention, and Outcomes within a Randomized Experiment*

*Presenter: Rebecca Thornton*

This paper evaluates a program that randomized incentives to obtain health insurance among informal sector workers in Managua, Nicaragua.
In January, 2006, the government of Nicaragua extended a health insurance program to informal sector workers using microfinance institutions. We randomly varied the costs of enrolling as well as randomly assigned different affiliation locations to sign up for the insurance: either at the central office, or the participating microfinance institutions. Costs were crucial to signing up for health insurance – both monetary costs as represented by the subsidies offered as well as convenience costs as represented by a direct marketing effort to sign individuals up at their place of work.

Approximately one year after being offered the insurance, insured individuals switched from using services at private and Ministry of Health facilities to visiting covered health facilities contracted by INSS; overall utilization of services did not increase. Total out-of-pocket expenditures were reduced among insured individuals, but the average out-of-pocket savings were lower than the equivalent unsubsidized insurance premiums. We also found very low retention rates after expiration of the subsidy, with less than 10 percent still enrolled in the insurance program after one year. In addition to the empirical results, we discuss the institutional challenges and limitations of this project that are essential to consider for other similar programs.

| Themes 2: Impact evaluation: Approaches and methods | Session number: [23.7] |
| Themes 6: Impact evaluation of policies, programs, projects, partnerships and networks | |
| Session stream: Health | Session name: Health |

Health Services Extension Program - Ethiopia

Presenter: Assefa Admassie

HSEP is a community based initiative for the provision of primary care to households. Operating at the village (or Kebele) level, the program involves a health post with two health extension workers providing preventive and curative services including maternal and child health care, immunization, information on nutrition, malaria prevention and control and health education to approximately 5000 residents each. The evaluation applies propensity score matching at the village level across three administrative regions. The sample size consists of 140 (72 treatment and 68 control) randomly selected rural villages with a random selection of 20-25 households in each village. Child specific health outcomes include: incidence of diarrhoea and duration, immunization rates for major diseases and incidence of coughs amongst children below five years. Maternal health outcomes analysed cover: TT2 vaccination during pregnancy, number of deliveries attended by skilled healthcare workers and usage of post-natal care utilization.
Evaluating an Intervention to Improve Maternal Health

Presenters: Juan José Díaz, Miguel Jaramillo

The health of mothers and children are two policy priorities in Peru for good reasons. Even after a decade of improvements in the Peruvian health sector (Cotlear 2000, Valdivia 2002) and consequent reductions in maternal and child mortality rates, these are still among the highest in the region. In effect, the rate of maternal mortality in Peru (185 per 100,000 births) is the third largest in Latin America and the Caribbean (LAC), more than doubling the average for the region, and well above that of countries with similar levels of development (DHS 2000). Further, there are very large differences between rates for poor and non-poor. If we look at factors associated with these rates, we find that use of both institutional delivery and prenatal care are low and inequitably distributed. While on average in urban areas 91.5% of births are professionally assisted, the same figure is only 43% in rural areas. Though health facilities sometimes do not ensure adequate treatment, case-fatality rates for home births are much higher. Recent research has also identified cultural/ethnic, gender-related as well as geographic barriers in access to maternal care (Jaramillo 2006). Within this context, the MDG concerning maternal health has provided additional pressure for the Peruvian government to act on this front.

The problem of high mortality rates certainly concerns most of the underdeveloped world, African and Asian countries as well, particularly Sub-Saharan Africa and south Asia, where the problem is most severe. The causes of maternal death are similar across Latin America, Africa, and Asia: post-partum severe bleeding, infections, hypersensitive diseases, and unsafe abortion. In sub-Saharan Africa, the combined maternal mortality ratio for severe bleeding, hypertensive diseases, and infections is around 500 deaths per 100 000 live births, compared with fewer than 300 per 100 000 in south Asia, just over 100 in Latin America and the Caribbean, and four per 100 000 in developed nations (Ronsmans and Graham, 2006). Within-country inequalities in access to health services are generally associated to these appalling numbers, the rural areas typically being the less favored.

We evaluate the Peruvian Health Reform Support Program (PARSalud), a public investment program oriented to improve maternal health. The program operates at eight regional offices of the Ministry of Health, selected because they serve the population with the highest maternal and child mortality rates in the country. The implementation of PARSalud is based on a “causality model” that attacks the sources of maternal mortality and morbidity by seeking to eliminate economic, access, and cultural barriers to modern health care use amongst the poor.

We use administrative data and exploit spatial and time variation in the execution of the program to implement difference-in-difference methods to identify the impact of the program at the facility level. We find a positive impact of the program on the number of attended births, number of births using the clinical protocol, and number of cesarean interventions. In a second paper we evaluate the impact of the program at the mother's level, focusing on both objective (compliance with clinical protocols) and subjective (mothers’ perception) indicators of quality of care at delivery. We identify impacts on the former, but not on the latter.
Behavior Change Interventions to Prevent HIV among Low-Income Girls and Women Living in Low and Middle Income Countries

**Presenter: Sandra McCoy**

This paper presents a systematic review of behaviour change interventions to prevent HIV. The review includes studies of behavior change interventions that target individuals as well as those targeting communities. It reviews a variety of strategies and tools, including: conditional and unconditional cash transfers; targeted messages and social marketing campaigns; school-based HIV prevention education; HIV testing; vocational training; empowerment training; non-cash incentives (like school uniforms); social support programs; and other interventions that have been subject or experimental or quasi-experimental evaluation. Cost-effectiveness of different interventions is analyzed.

Randomized Experiments Evaluating Effectiveness of HIV Prevention Programs

**Presenter: Rebecca Thornton**

Over the past two decades, the HIV/AIDS epidemic has afflicted millions of individuals in Africa. In the absence of significantly expanded prevention and treatment programs, the epidemic is expected to worsen in many other parts of the world. Since the beginning of the HIV/AIDS epidemic, various strategies and interventions have been implemented in order to curb the spread of the disease and prevent further infections. This presentation presents results from several randomized evaluations of several of these policies. There are three experiments. Each experiment was conducted in Malawi, between 2004 and 2008 and include between 1,400 and 3,000 respondents throughout the country and include:

- HIV testing to know one’s own or one’s partner’s status (VCT)

This experiment randomly assigned monetary incentives to individuals to learn their HIV results after being tested. Distance to the HIV results centers was also randomly assigned. Without any incentive, 34 percent of the participants learned their HIV results. However, even the smallest incentive doubled that share. Using the randomly assigned incentives and distance from results centers as instruments for the knowledge of HIV status, sexually active HIV-positives who learned their results are three times more likely to purchase condoms two months later, than HIV-positives who did not learn their results; however, they purchase only two additional condoms. There is no significant effect of learning HIV-negative status on purchasing condoms.
• Providing randomized cash incentives to stay HIV negative.

This paper evaluates a new HIV prevention strategy: offering conditional financial incentives for individuals (and couples) to maintain their HIV status. Approximately 1340 individuals in rural Malawi were tested for HIV in 2006. They were then offered financial incentives of random amounts ranging from zero to values worth approximately four month’s wage if they maintained their HIV status for 15 months. Three times after being offered the financial incentives at the beginning of the study, respondents were asked about their sexual behavior through interviewer-administered sexual diaries. Our preliminary findings suggest no behavioral change in response to the financial incentives. However, we find some suggestive evidence of effects of the incentives among several sub-groups: those with higher annual incomes and women with higher levels of empowerment who were offered a financial incentive were more likely to report having condoms at home. While these preliminary results may shed light on some of the barriers to HIV prevention behavior, there are no estimated effects on the likelihood of reporting using condoms during sexual activity or on the likelihood of engaging in vaginal sex.

• Providing information about the relationship between male circumcision and HIV transmission This study assess whether the dissemination of information that circumcision reduces the probability of HIV transmission affects (i) the demand for circumcision among uncircumcised men and (ii) the demand for safe sex among both uncircumcised and already circumcised men. A baseline study was conducted in 2008 in which villages were randomly allocated information that circumcision is protective against HIV transmission.

| Themes 2: Impact evaluation: Approaches and methods | Session number: [24] |
| Themes 6: Impact evaluation of policies, programs, projects, partnerships and networks | |
| Session stream: HIV/AIDS | Session name: HIV/AIDS |

Indicators for assessing the impact of new funding modalities on civil society responses to HIV/AIDS

**Presenter: Kevin Kelly**

The ‘Three Ones’ (an international agreement for structuring national HIV/AIDS responses) and the ‘Paris Declaration on aid effectiveness’, have reshaped international funding to support HIV/AIDS responses. There has been relatively little attention paid to the impact of the new funding modalities – including general budget support, joint funding arrangements and basket funds – on civil society agencies at the forefront of community HIV/AIDS responses. This needs to be examined, as many civil society organisations have traditionally relied on direct support from international donors and this has become increasingly difficult to access.

The paper presents an overview of changes in funding support for civil society, involving interviews with donors, government HIV/AIDS agencies and civil society leaders in nine countries. In addition a survey of 439 civil society organisations was conducted in six countries, and case studies were conducted of local responses to HIV/AIDS and the support they receive. It was shown that civil society responses to HIV/AIDS have increasingly been supported by national HIV/AIDS sub-granting mechanisms and direct support from donors is waning. In some countries much of the support that remains is provided through the conduit of international NGOs acting as coordinating agencies within national HIV/AIDS architecture. This poses significant risks to civil society organisation development, diversity and responsiveness. The new funding approaches may limit the contribution of national civil
society organisations and ultimately the unique contribution that they make to national HIV/AIDS responses. There is a need for donors and national civil society organisations to take note of the risks to civil society responses in the current funding environment, and to re-engage in supporting the development of strong national NGOs and umbrella bodies.

The paper suggests a number of indicators, derived through this research, which may prove useful for understanding the impact of the new funding modalities on the evolution and ultimately effectiveness of civil society organisations working in the HIV/AIDS field.

### Themes 2: Impact evaluation: Approaches and methods

### Themes 6: Impact evaluation of policies, programs, projects, partnerships and networks

**Session number: [24]**

**Session stream: HIV/AIDS**

**Session name: HIV/AIDS**

**Planned and unexpected outcomes and impact of HIV peer education and mitigation in different Southern African communities.**

**Presenters: Madri Jansen van Rensburg, Amorlett Nyamweda**

The HIV epidemic is affecting Southern African communities with most of the people living with HIV and AIDS found in Southern Africa. Many HIV prevention projects use peer education as a participatory method across age groups and in different settings. Projects directly target high risk groups (e.g. sex workers) or groups with high prevalence rates such as the youth or are broad community based. Community volunteers are used as caregivers and care supporters in HIV mitigation projects.

Regional and intermediary HIV organisations play a major role in supporting grassroots organisations. They also provide support for monitoring and evaluation efforts, including capacity building and assisting with outcomes and impact studies. Measuring the outcomes and impact of projects are important. It is also crucial to measure the regional impact (and cross-country impact) of interventions. Describing the changes in evaluation tools and methodologies are important for regional organisations to do and influence evaluations.

This study aims to describe the planned (expected) and unexpected outcomes and impact of HIV peer education prevention and mitigation projects. It further aims to describe the relationship between documented unexpected outcomes and different types of methodologies used (including quantitative, qualitative and mixed methods). It further investigates the processes and contributing factors that influence outcomes and the variation in the outcomes.

The study includes a review of more than thirty evaluation studies from two regional Southern African HIV organisations. It includes studies from Botswana, Lesotho, Malawi, Mozambique, South Africa, Swaziland, Zambia and Zimbabwe. There is a focus on cross-country and cross-cultural analysis. Studies for review include quantitative and qualitative and mixed method approaches. The studies include community based projects and other settings such as the workplace and tertiary education institutions.

All the studies included in the review investigated long term projects with clear objectives and targets. All projects included participatory methods and interventions. The HIV prevention projects were based on peer education and the HIV mitigation projects on community volunteers as caregivers and care supporters. The review employs a mixed method approach. It includes statistical analysis of the available quantitative data and qualitative analysis of information gathered from open-ended questions and qualitative data from focus group
discussions, in-depth interviews, Most Significant Change stories and other qualitative methods. It includes impact findings of process evaluations.

In many instances unexpected outcomes are reported. This includes developmental changes in communities beyond HIV prevention. Findings also include unexpected beneficiaries beyond the initial target groups. Methodologies of some longitudinal studies were changed to include the capture of unexpected results (e.g. mixed method approaches), the expansion of tools to investigate relevant issues (e.g. male circumcision) and capacity development of researchers. These changes had to include cross-cultural appropriateness and validity. This review has practical implications for regional organisations planning evaluation studies, developing monitoring tools and managing monitoring and evaluation systems.

### How effective are social programs during conflicts? Instrumental variables evidence from the Angolan civil war

**Presenter: Eric Djimeu Wouabe**

In this paper, we consider the impact on child anthropometrics and household expenditures of the only major social program in Angola during the civil war, the World Bank sponsored Fundo de Apoio Social (FAS). Our identification strategy is based on Angolan politics at the time of program deployment. Given that the FAS was meant to contribute to the process of national reconciliation following the Lusaka Protocols that brought a temporary lull in the fighting, we use the distance to the main UNITA rebel bases of the period to generate plausibly exogenous variation in treatment status. Our linear instrumental variables estimates show that treatment by the FAS during the 1994-2000 period was associated with a 48.5% increase in household expenditures per adult equivalent, and a 34.5% increase in child height-for-age. We then use the local instrumental variables (LIV) estimator to explore the marginal treatment effect (MTE) of the FAS. In doing so, we find that each 1000 additional deaths attributable to the civil war within a 5 km radius of the community shifts the MTE associated with household income per adult equivalent up by 7.1%. The FAS was therefore associated with substantial benefits for treated communities, with these benefits being significantly increasing in the intensity of violence faced by the inhabitants.
The (Potential) Impact of Public Works Programmes on the Lives of the Poor and the Vulnerable in a Post-Conflict Environment

Presenter: Doreen Chemutai

Put into perspective, despite the impressive poverty reduction recorded nationally in Uganda, the North has remained the poorest region, with income poverty nearly two-fold that of Uganda. The two-decade long civil conflict and low human development are the commonly cited explanations for the poor progress in poverty reduction in this part of Uganda. The government and other development players are committed towards improving the livelihoods of the people in Northern Uganda, with particular focus on the poor and vulnerable. One of the interventions that is popularly employed by government and non-government service providers is public works programmes (PWPs).

While there are variations in the mode of implementing PWPs across different players, the common thread among PWPs is usually the purpose: providing opportunities for cash-income through sale of unskilled labour, and simultaneously creating / maintaining community infrastructure. In this way, the proponents contend that both income-based and access-based avenues will be employed to improve the standard of living of the vulnerable and poor in Northern Uganda.

The two key evaluative questions to investigate the potential impacts of PWPs are:

To what extent are public works an effective strategy of poverty reduction among the vulnerable - breaking the vicious cycle of poverty; regaining human dignity and hope in life; bringing back their self-esteem and confidence to participate actively in communal activities and live a “respectful” and dignified life? How effective are they for those with “chronic incapacity to work and earn”?

To what extent is consumption smoothing and equity among communities realised through re-distribution of income (transferring cash-incomes to the vulnerable)?

One of the key assumptions that underpins many PWPs and will be investigated by the study is whether PWPs are truly self-targeting: that the interventions actually reach the most vulnerable households; that a significant majority of vulnerable households have at least an able-bodied person who is not only willing to: a) earn bread for the family; b) exchange labour for a low wage; and c) but also has the time to engage in the public works

The study will focus on one case study – a PWP implemented by one of the development actors in Northern Uganda. Preference will be given to a PWP with multiple sites spread across at least 3 districts (Acholi sub-region) of Northern Uganda.
The global evaluation of Rewrite the Future Campaign in 4 conflict-affected fragile countries with key focus on the process, findings and lessons learned from Southern Sudan

**Presenter: Dragana Sretenov**

The paper will summarise the mid-term learning and findings from the first global level evaluation of Rewrite the Future Campaign (RtF) conducted by the Save the Children Alliance in 4 conflict-affected fragile countries.

Of the 77 million children out of school today, 39 million—more than half—live in conflict-affected fragile states. In many of these countries years of instability and conflict have devastated the education system. Even when children living in these countries do have the chance to go to school, the quality of education they receive is frequently substandard and they are less likely to complete even a basic education.

To address these issues, Save the Children has pioneered a global campaign: *Rewrite the Future*. In 2005, Save the Children initiated education programmes in 20 countries around the world to serve the needs of the children affected by armed conflict. The overall global goal of *Rewrite the Future* is that: *All children affected by armed conflict have an opportunity to fulfil their right to an education that enables them to learn, play and develop*.

The purpose of this longitudinal evaluation of RtF is to identify effective ways of delivering quality education in conflict-affected fragile countries, and contribute to the documentation and learning on this subject. Donors and external community of practice are the main target audience of the evaluation.

The evaluation is focusing on education quality, measured through aspects of learning, relevancy, participation and a safe learning environment. All countries are exploring the two global research questions:

1. How have Save the Children’s *project level* interventions contributed to *quality primary education* for children affected by conflict?
2. Which project level interventions have had what *impact* on the education quality of children affected by conflict?

The methodology is qualitative, quantitative and participatory and in each country the evaluation teams led by the global researcher have selected a specific quality intervention for in depth impact analyses. The recommended global methodology is a truncated longitudinal design, midterm and post test, with a comparison group.

The evaluation is taking place in 4 countries: Nepal, Afghanistan, Angola, and Southern Sudan and the teams will perform the two data collection processes, and the first evaluation phase has been successfully completed in November 2008. A mid-term report of findings will be produced by mid March and shared internally and externally. The final research report will be available in early 2011.
The presentation will focus on the findings from the mid-term report and will cover: Global planning; Methodology and data collection plan; Country implementation; Preliminary findings; and Lessons Learned in all 4 countries with specific focus on Sothern Sudan.

Using innovation to document project impact in volatile situations

**Presenter: Haron Njiru**

The need to know how any health intervention is helping alleviate a health problem is a critical component in health programming. However this is not always possible as evidenced by the experience of Merlin in the Somalia republic.

Since 2005, Merlin has undertaken disease surveillance and outbreak response activities in Puntland as part of communicable diseases prevention and control following the Tsunami that affected the coastal strip of Puntland Somalia. The strategy of this project is to strengthen disease surveillance, emergency preparedness and effective response to outbreaks with an aim of minimizing mortality and morbidity attributable to 5 epidemic prone diseases in the area.

Despite there being a desire to monitor the progress of the intervention, it this was not practical for several reasons. Among them, there was the fact that the project areas were spread over a large area, the data capture tools were not uniform hence one could not gather the data into a central template for comparison and analysis, and the most distressing fact was the volatile political climate which made it impossible to regularly travel to the data generation points for data collection. It was generally not possible to neither know how the project was doing nor justify the continued funding of the intervention.

Late, 2007 a consultant was hired to assist in addressing this issue. Among the key recommendations by the consultant was to harmonize the data fields, to computerize the process of data collection and start electronic reporting. This way a stream of uniform data could be sent to a central data repository for archiving and analysis. This would enable the project impact to be documented even when there roads are blocked by militias.

This use of information technology has helped Merlin overcome the barrier to evaluation, increase the efficiency of the project and reduce the time spent in cleaning data. Countries experiencing political instability can make use of information technology approach to bypass some of the barriers in impact documentation process. This would keep the donors informed and ensure continued donor support to the communities in need of humanitarian or other assistance.
Impact Evaluation for Livelihood Creations in Post conflict Situation

*Presenter: A. Ariyasuthan*

This idea of “livelihoods” is not a new concept. A livelihood is what all of us have to enable us to live our lives. How successful our livelihood is in helping us reach the goal of living healthy and balanced lives is what we are concerning ourselves with in this manual.

While projects typically target only a portion of families within a community, thought should be given as to how these projects will also encourage growth or coping mechanisms in the community as a whole. When making decisions about who the beneficiary is, what the type of project is, and if any additional “value added” projects should be implemented (such as a revolving loan fund within a CBO), the overall community needs to be examined.

The livelihood outcomes that appear here are may or may not be relevant in any given situation. What determines what needs to happen in a community will be up to the community members, and therefore may include some, all, or even additional outcomes than what is listed below. It should be recognized that at the core of all the outcomes below, an ability to provide for the family and maintain security in terms of food, health and education is the central factor. This is not to say that "money provides happiness", but rather to emphasize that the only way for the poor to reduce their vulnerability and maintain an even basic level of nourishment and health is through the generation of resources.

Because our projects are funded by outside money, it is critical that we are able to show results. Based on survey and assessment information, we should be able to show improvement in key areas. This should be quantifiable. So for example, if we are implementing a livestock project, it is not enough to say that the project was successful because the livestock did not die. How did the livestock project specifically improve the lives of the beneficiaries? Was there an increase in income? Did it provide the family with a financial insurance, knowing that they could sell an animal if an emergency came up? Did it improve their food security? This is a very important aspect from the reporting side. It also tells us if the projects we are implementing are having the best impact, and if not, that maybe we should consider other types of projects.
Experience of monitoring education in post conflict Rwanda

Presenter: Thaddee Yossa

De 2003 à 2006, le Ministère de l’Education (MINEDUC) de la République du Rwanda et InWEnt Capacity Building International, Allemagne, ont mis en œuvre le projet «Formation en pédagogie active pour renforcer une culture de paix à l’école », dont la finalité est de contribuer à la construction d'une société démocratique, capable de gérer et résoudre ses conflits par des moyens pacifiques et non violents et de surmonter les traumatismes de la guerre, du génocide et de la violence.

Ce projet visait à introduire et développer une pédagogie active, des méthodes participatives et les compétences en gestion des conflits comme orientation transversale de la formation des enseignants pour renforcer une culture de paix à l’école. Plus spécifiquement il s’agissait de renforcer les capacités pédagogiques des enseignants des écoles normales des instituteurs et du primaire et d’améliorer l’environnement d’apprentissage dans les écoles.

Le projet a formé au niveau national 8 multiplicateurs, 27 concepteurs de programmes, 10 inspecteurs pédagogiques, et dans la province de Kibungo 10 animateurs pédagogiques, 34 enseignants de TTC, 2 inspecteurs de district, 2 représentants d’association de parents d’élèves et enseignants, près de 10 000 élèves de 14 écoles primaires, 125 enseignants d’écoles primaires, 14 directeurs d’écoles primaires.

Le système de suivi évaluation élaboré et mis en œuvre visait à capter les effets et impacts escomptés relativement aux aspects suivants:

- Renforcement de la pratique de la pédagogie active dans les écoles primaires/TTC
- Amélioration du climat d’apprentissage dans les écoles primaires/TTC
- Amélioration de la qualité des accompagnements pédagogiques dans les écoles
- Développement renforcement de la sensibilité Genre et respect des différences dans les écoles et la communauté
- Développement renforcement Capacité de gestion pacifique des conflits et prévalence de l’état de droit sur l’état de fait dans les écoles et la communauté
- Meilleur épanouissement des élèves : créativité et autonomie accrues
- Meilleur respect de la vie et de la dignité humaine dans les écoles
- Renforcement des capacités de préservation de l’environnement.

Les difficultés à cerner et capter les impacts du projet et la façon dont ces difficultés ont été gérées peuvent être intéressantes à partager.
Approaches to Institutionalizing IE: the Spanish Impact Evaluation fund and the Millennium Challenge Corporation Models

**Presenters: Sebastian Martinez, Frank Wiebe**

This session will present two different approaches to institutionalizing impact evaluation in large development donor agencies. The two approaches differ in the ways they address the need for development institutions to do more impact evaluations, and incorporate examples of how to manage and commission impact evaluations and use the results to improve organizational policy. The approaches also illustrate ways of strengthening organizational cultures, values and resources through training, support and stimulation of demand.

**Presentation 1**

**Title:** The Spanish Impact Evaluation’s Fund (SIEF) model of institutionalizing Impact Evaluation of World Bank Funded Human Development Projects

**Presenter:** Sebastian Martinez, Economist, Human Development Network, World Bank

The Human Development Network at the World Bank encompasses the World Bank’s work in health, education and social protection. The Spanish Impact Evaluation Fund (SIEF) has aided in the institutionalization of impact evaluation not only in the HD Network, but throughout the World Bank drawing on the World Bank’s position as the premier global development institution to develop lessons based on cross-country evidence of ‘what works’ by comparing types of interventions across countries. With the support of SIEF, Impact evaluations are organized into clusters based on the types of interventions; this approach enables researchers to compare the effectiveness of individual interventions in disparate conditions and across alternative designs. There are six main research clusters at present: HIV/AIDS Prevention; Conditional Cash Transfers; Basic Education Accountability; Malaria Control; Health Contracting/Performance; and Active Labor Markets/Youth Employment. Efforts are made to conduct evaluations jointly with governments, helping build capacity to manage for results and ensure the relevance of the research to policy. SIEF currently funds 47 impact evaluations in 33 different countries.

In addition to supporting impact evaluations financially, part of the SIEF’s priority agenda includes conducting 5 regional workshops per year designed to provide customized training in impact evaluation. Participants are represented by teams of government officials, local researchers and World Bank task teams who are work jointly together on the design of new World Bank-supported projects. The goal is to equip these teams with technical skills needed to build high quality impact evaluation into project design.

Consistent and rigorous evaluation of new government programs will build an evidence base for better HD policy that will benefit countries and enhance aid effectiveness. One of the best indicators of SIEF’s success is demonstrated by continued regional and national interest in follow-up workshops. These workshops also serve as a catalyst whereby workshop participants are able to get a head start on designing and operationalizing their specific evaluations.

**Presentation 2**

**Presentation Title:** Focusing on Results: How MCC Uses Impact Evaluation to Promote Aid Effectiveness
Presenter: Franck Wiebe, Chief Economist, MCC

The Millennium Challenge Corporation’s (MCC) approach to aid reflects principles that now are widely accepted in the development community: economic growth matters; good governance matters; country ownership matters; and results matter. MCC uses the term “impact evaluation” only in reference to the most rigorous, quantitative assessments that include an explicit formulation of the counterfactual. These IEs are contracted to independent third party evaluation firms, many connected to universities. MCC funds IEs for roughly 15% of all MCC program activities. For all other activities, however, MCC requires a formal assessment of program performance, and this will include a comparison on benchmarks established in the benefit-cost analysis that established the economic rationale of the investment.

MCC has instituted its impact evaluations within a larger “results framework” that follows every project from conception to completion. Once a project has been proposed for MCC assistance, MCC conducts a benefit-cost analysis to estimate the likely impact of the project. Many of the key parameters from the benefit-cost analysis become indicators in the monitoring and evaluation plan for each project. For a subset of project activities, MCC hires independent organizations to conduct rigorous impact evaluations. MCC uses the results and data gathered from impact evaluations to inform future project designs and pre-investment analyses. These evaluations also contribute to the body of “best practices” in the broader development community. MCC posts information about its impact evaluations on its website. This initiative is intended to encourage broader technical exchange within the impact evaluation field.

Table: Themes 3: Institutionalizing impact evaluation: demand, production and use

<table>
<thead>
<tr>
<th>Session stream: Institutionalizing impact evaluation</th>
<th>Session name: Institutionalizing impact evaluation</th>
</tr>
</thead>
</table>

Institutionalizing Impact Evaluation: A case study of Sri Lanka

Presenter: Sarojini Dias

Decentralizing the highly centralized system of education management was attempted at seven occasions in Sri Lanka from 1961-2006. Programme for School Improvement (PSI), with the aim of giving more autonomy to the school for quality improvement is the latest. To identify the impact of PSI on school management the present study was designed in three phases.

In the first phase (2006), the state of the art of school management was studied as a baseline survey (BLS) on 6 themes related to school standards (School level planning, Role of the Principal, Curricular and Co-curricular activities, Teacher and Student development, School community relations, Problems and challenges). 35 schools exposed to PSI and 35 schools not exposed were the sample. Using six interview schedules for principals, teachers, students, parents, community members and education officials 70 school profiles were prepared under standards and comparisons were made between PSI and non PSI schools. There was no significant difference between two categories. 99% of the schools were equipped with necessary planning documents, lack of awareness and training in school level planning was a major factor impeding the implementation of the plans.

In the second phase (2007), in-depth case studies were conducted on 07 PSI schools selected based on the features identified during BLS:

- Situated in a remote area with problems in accessibility
- Energetic leadership
- Trying to rebuilt after a disaster
• Constant transfers of principals
• No data base for the school
• Situated in a metropolitan area but low demand for school
• Political interference for school admissions

The case studies enabled to clarify issues and to identify specific needs of schools to make the PSI effective. Findings of the BLS and the case studies were presented to relevant authorities and won their confidence and support for further research. Arrangements have been made to start the impact evaluation phase in 2009 using the same sample, standards and data collectors. Interview schedules were refined to rate the position of each standards across a five point scale and to record qualitative data. The effect of PSI will be assessed by comparing the BLS situation with the situation after 3 years of its implementation.

The impact evaluation is a new venture, as research based information on baseline status and impact of interventions are dearth in Sri Lankan research and education culture resulting drawbacks to the reform efforts. As a result, the design and the collaborative nature of the study have created high demand for its results and the findings will be disseminated through print and electronic media and through seminars in December 2009 to institutionalize impact evaluation at different strata of education system including ministry and school community.

Developing the capacity on research culture is a necessary factor in institutionalizing impact evaluations. Data collectors were given an extensive training, giving the opportunity to discuss and clarify data collection procedures. To build capacity of researchers and data collectors on both quantitative and qualitative research, however, there is a necessity for more orientation and experience on research.

<table>
<thead>
<tr>
<th>Themes 3: Institutionalizing impact evaluation: demand, production and use</th>
<th>Session number: [26.2]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session stream: Institutionalizing impact evaluation</td>
<td>Session name: Institutionalizing impact evaluation</td>
</tr>
</tbody>
</table>

Challenges to Institutionalizing Impact Evaluation – Illustrated by joint government donor supported Monitoring and Evaluation of a Food Security Program in Ethiopia

**Presenter: Getahun Tafesse**

This paper aims to contribute in the effort of institutionalizing evaluation by addressing challenges and possible remedial actions. Its specific objectives include the following:

1. briefly review the theoretic foundations of evaluation
2. synthesize existing global and national context for evaluation
3. elaborate challenges faced in institutionalizing evaluation
4. share useful experiences from a case of a multi-donor and billion dollar financed food security program in Ethiopia
5. identify key recommendations to strengthen institutionalized evaluation

The paper will review the experiences of donors and government in Ethiopia in strengthening evaluation activities around a key multi-donor financed food security program in the country. The review will address government implementation capacity at different levels, organisational culture that values and supports evaluation and possible approaches for organizational learning and informed decision making. It will also share useful experiences based on a case of a multi-donor billion dollar financed food security program in Ethiopia.
Assessing programme effectiveness at the global level in a large and complex organisation

Presenters: Irko Zuurmond, Junaid Habib

How does a large and complex organisation know whether globally its programmes are effective? How, at the global level, can it optimize the use of information that is generated by the myriad of programme initiatives in multiple countries, to generate an informed assessment of its programme effectiveness? These are some of the questions that organizations like Plan are struggling with.

Plan is a large development organization with programmes in 49 countries. Plan’s programmes cover a wide range of different sectors including health, water and sanitation, education, economic security, child protection, etc. The programme focus in each country depends on the identification of the issues that most affect children as well as Plan’s niche in country.

In 2003 Plan adopted Child Centred Community Development (CCCD) as its overarching approach to development. CCCD is a rights-based approach in which Plan is a facilitator of development processes that contribute to the fulfilment of child rights as defined in the Convention on the Rights of the Child and other rights instruments. To this effect Plan works with children, families, communities, civil society organisations and the government at different levels in recognition of their different roles and responsibilities vis-à-vis the realization of child rights. While the programme focus differs from country to country, CCCD is the common denominator that is reflected in all programmes regardless of the context in which they are implemented.

At the level of programme countries there have been clear corporate processes in place for planning, monitoring and evaluation since 1997. Despite the wealth of data and information produced through these PME processes at the level of programme countries, making an informed assessment of Plan’s programme effectiveness at the global level remains a challenge. To address this situation Plan has developed an overarching Programme Effectiveness Framework.

To ensure a consistent assessment of programme effectiveness the framework proposes levels and dimensions of change as the analytical lens applied across the programme initiatives in Plan’s 49 programme countries. This lens has been developed by unpacking CCCD as Plan’s overarching intervention logic. Consequently it focuses on change at the level of rights holders, duty bearers and civil society organisations at different levels. For each type of social actor, the framework identifies those outcomes that are considered essential in bringing about lasting change at impact level.

At the global level the framework proposes a focus on assessing the effectiveness of strategies and policies through a mix of initiatives, using multiple methodologies and multiple sources of information, in recognition of the complexity of social, cultural and political change. These include, amongst others: analysis of data on 1.5 million sponsored children; trends in output data; multi-country thematic evaluations; post-intervention studies; annual meta-analysis of programme progress and evaluation reports. These different analyses culminate in the compilation of a three-yearly Global Effectiveness Report which synthesizes and triangulates the effectiveness information generated by different parts of the organization. This will enable Plan to make an informed assessment of programme effectiveness at the global level.
Evaluating the level of institutionalization of Participatory Planning Monitoring & Evaluation (PPM&E) in the National Agricultural Research System in Tanzania

Presenter: Cypridion C. Mushongi

In order to improve the impact of agricultural research investment and relevance of agricultural research, the system, procedures and methods have been developed and thence become integral part of client oriented research management system. In effecting this, the agricultural research institutions within the Ministry of Agriculture, Food Security & Cooperative have recognized the importance of PPM&E and its the potential contribution in enhancing the effectiveness of the technology development process. As such, PPM&E of research activities is integrated into the technology development cycle, from problem identification to dissemination of recommendations and impact assessment. Nevertheless, agricultural research institutions have been working to develop evaluation capacity, with a view of contributing to impact and sustainability of agricultural research interventions. However, under budgetary pressure and demand from the agro-industrial sector most of National Agricultural Research Institutions in Tanzania and other developing countries face intense financial crises. Furthermore, the effects of globalization and trade liberalization, and the need to contribute to poverty alleviation as indicated in the MDG’s have further aggravated the challenges for research institutions.

Thus, in order to face these challenges research institutions have increasingly pursued options for decentralization, privatization, as well as institutionalization of PPM&E so as to increase client participation in technology development. However, in relation to institutionalizing PPM&E, three main questions are not yet answered since the advent of PPM&E: firstly, are PPM&E systems functioning? Secondly, to what extent have they been institutionalized? And thirdly, are PPM&E systems operating systematically, effectively and efficiently? The answer for the questions above would help to address the key challenge in monitoring and evaluation of agricultural research institutions, which is to develop ownership and get feedback from clients and stakeholders on whether activities contribute to the planned output of the research programme, while using inputs by all parties as planned.

Hence therefore, this proposal envisage to determine the extent of institutionalization of PPM&E systems and measuring the performance of their functions, taking into account of the approach developed within the framework of Client Oriented Research Management Approach (CORMA) over the past 8 years. The method to be used to evaluate PPM&E systems will be based on the conceptual framework, which will be developed from three dimension namely: (a) an appropriate operating framework; (b) a set of quality research products designed to provide feedback on the intervention strategy of operations; and (c) a strategy for use and communication of research outputs/results with institutionalized forums so that actors can use the information within social learning processes.
We shall learn, but shall we use

**Presenters: Jocelyne Delarue, Jean-David Naudet, Veronique Sauvat**

Impact evaluations of development interventions constitute a unique opportunity to bridge the gap between researchers and policy makers, and contribute to development effectiveness. Yet there is no guarantee that this knowledge will translate into more effective programs. Indeed, while drug trials are routinely used in the pharmaceutical industry to favor or reject products and study their side effects, the experience in the domain of social interventions remains disappointing: if a few impact evaluations have made their way into public policies, a much larger number have remained unused by policy makers. Understanding the underlying reasons can contribute to avoid excessive expectations from such evaluations while enhancing their actual use. Drawing on a review of the evaluation literature in more advanced countries and an on-going evaluation of a health insurance project in Cambodia, this paper aims to derive lessons relevant to the field of development.

We distinguish between *instrumental use* whereby an evaluation directly influences program’s or policy’s design; *conceptual use* whereby it indirectly changes policy makers ideas by contributing to the overall knowledge in a given field; and *persuasive use* where policy communities or issue networks selectively use evaluations results that best support their advocacy work. We find that instrumental use of evaluations only happens punctually and under favorable conditions (limited changes in policies, strong interactions between policy makers and evaluators, timeliness with respect to a policy window). Although mostly untraceable, conceptual use is probably more frequent, and likely enhanced by robust – and therefore comparable – impact evaluations which facilitate knowledge accumulation. Recent efforts towards meta-evaluations further increase this potential by addressing external validity critics often raised by development specialists. Finally, persuasive use can be significant and limit the knowledge generating process. This will be particularly the case for donor-funded evaluations aimed at demonstrating that aid works rather than understanding what works better. It is further complemented by a tendency to report the positive and/or surprising results more than non-significant ones – a bias well identified by initiatives such as 3ie.

Overall, use of impact evaluations can be significantly enhanced through (i) better match of knowledge supply and demand at a given moment, and (ii) better appropriation of results by policy-makers. Because impact evaluations may involve various interests from managers, donors and researchers, compromises are necessary to ensure full impact of such studies on development outcomes. Under certain conditions, interfaces such as evaluation units – in donor agencies but also ministries from aid recipient countries – can help in the matching process.
We know we need to do it, we really want to do it, but - we never do because...

**Presenter: Anzel Schonfeldt**

In the last decade, South Africa has seen an influx of billions of rands of donor funding to address a wide range of health, development and social issues from donor powerhouses such as the US Government, the Global Fund for AIDS, Tuberculosis and Malaria, The Gates Foundation, The Ford Foundation and the EU. Continual funding is often dependent on the monitoring of programmatic results against targets. Only recently has donors begun to move towards impact evaluation as a programmatic/project outcome, although relatively little funds are forthcoming to meet these evaluation needs. Although donors argue that the bulk of any funding should be utilized to reach intended beneficiaries, a lack of funds to do proper impact evaluation may in fact hurt the very communities they are trying to assist in the long run.

There is no denying that we know how important impact evaluation is. Most donor-funded organization’s monitoring and evaluation staff members in South Africa have attended at least two workshops on the subject, can easily rattle off at least five major benefits of impact evaluation and can readily produce several work plans detailing intricate timelines; identify often spot on methodologies and harbor the best of intentions to disseminate impact evaluation results accordingly.

Why do we rarely evaluate impact at NGO or donor recipient level? Impact evaluation is expensive and time consuming. With pressure to implement and monitor, evaluation has fallen by the wayside, often nothing more than an afterthought to be sub-contracted to a consultant if the budget allows. Continuously changing indicators and indicator definitions, poor integration of indicators and targets between different donors, and ever-moving goalposts in terms of technical guidelines; the resulting training and retraining of staff to handle basing donor requirements (which often excludes evaluation) detract from any possibility to implement proper impact evaluation.

This paper proposes to examine the main reasons why South African NGO’s are currently producing very little impact evaluations in relation to the funding they receive from both local and international sources by means of in-depth interviews and surveys. Possible solutions will be explored and recommendations and best practices will be shared on how to promote impact evaluation in what can only be termed as a "hostile" environment where "M" disproportionally seems to forever come before "E" – regardless of the ultimate cost.
Evaluation and power: how can community organisation evaluation avoid un-developing?

**Presenter: Tracey Konstant**

Alongside local government, community-based organisations (CBOs) are the only organisational fabric with the breadth and depth of access to sufficient people, through which development might have potential for national or global impact. The Paris Declaration on Aid Effectiveness challenges international donors to support partner countries’ management systems. The great majority of evaluation systems for community organisations use systems that are owned and designed by the developed. This is to be expected. Few CBOs have their own evaluation systems.

Accepted systems tend to use due diligence criteria, programme indicators and targets, evidence-based planning, and similar. The present several shortcomings. Systems tend to permit those with high levels of literacy and abilities for sophisticated written communication to represent themselves far more convincingly than organisations which may be highly competent development practitioners, but have less formal education. Secondly, evaluation tends to rely on criteria for success selected by the donor or developed partner. Thirdly, they tend to be based in prediction of results, subjugating the real impact of emergent and responsive development. This use of processes designed at the convenience of the developed partner, with their control over criteria, place power firmly in the hands of the developed.

These externally designed, motivated and driven evaluation processes have the potential for profound negative impact on development effectiveness. Those systems which find their power in prediction, linear constructs and financial authority, can be particularly detrimental. Financially linked power in the hands of the developed is the source of global inequality and the fundamental tension that unpins the global development challenge of today. It is a power imbalance that can be reinforced by the exclusive use of approaches to evaluation that use the paradigms, criteria and form of the powerful.

This poses an interesting dilemma. Partnership choices have to be made. Financial decisions are unavoidable in dispersing development funding. Where the front-line development practitioner has little influence over the criteria or process by which they are judged, the very act of making these judgements is in conflict with development.

While a perfect solution is unlikely, and a variety of approaches should be used in communicating among partners, this paper considers some principles and approaches which might address this dilemma. Approaches that contrast with accepted modes of logic-based, often quantitative and pre-defined, written communication and unilaterally developed criteria for impact are explored. These ideas have been developed through a research process with HIV and AIDS Support CBOs in South Africa. The aim of the study is to design processes by which convincing outcomes and learning can be identified and communicated, with less emphasis on predicted results and linear arguments, and less requirement for high levels of literacy in order to fairly represent one’s contribution. We have attempted and observed methods, based on organisation-centred success criteria and formats, which support effective communication of the abilities and achievements of community organisation.
**Institutionalizing large scale experiments: Lessons from a recent evaluation of an unemployment counseling scheme in France**

**Presenter: Luc Behaghel**

Only very few large scale experiments have been carried out outside of the United States to evaluate government or other social programs. Because of their scale, these experiments raise a unique set of institutional issues and place special demand on stakeholder involvement that often can be ignored in small-scale evaluations.

During the course of 2007, more than 200,000 jobseekers became part of a large scale experiment to evaluate the effectiveness of two job counseling programs (one public and one privately run program) in France. In the experiment, persons at risks of long-term unemployment were randomly assigned into one of three groups: (1) intensive counseling by a government-run counseling scheme, (2) intensive counseling by a privately-run counseling scheme, or (3) the normal (less intensive) government track, and tracked for up to six months.

The evaluation was carried out across several regions in France and formed part of an ongoing reform of the French unemployment policy under the Plan d’Aide au Retour à l’Emploi (PARE). Preliminary findings suggest that the intensive government program significantly outperformed the privately operated program and the less intensive government track. The evaluation is ongoing and will also study the cost effectiveness of the unemployment counseling schemes and the long-term sustainability of employment for program participants.

The presentation will highlight the experiences of this large-scale experiment and key lessons learned related to the institutionalization of large scale experiments and the role of stakeholder involvement. Though these experiences are embedded in the French context, many lessons have implications for (and should be of interests to) policy-makers and researchers intending to conduct large scale experiments in development countries.

**Evaluating the impact of mass media education interventions**

**Presenter: Kevin Kelly**

The paper poses two challenges involved in evaluating the impact of mass media education interventions: 1) the problem of non-equivalent exposed and unexposed groups - given that the variables that influence exposure may also be associated with outcome variables; 2) challenges in measuring meaningful intervention exposure given that simple measures of exposure may not be sensitive to the theory of change on which the intervention is based. These problems are explored in describing the evaluation of behavioural outcomes of a mass-media HIV/AIDS education-entertainment intervention in South Africa. The evaluation methodology included a
960 person panel design and propensity score analysis. The main principles and processes involved in this method are described and it is suggested that this is an appropriate and sensitive method for evaluating behaviour change associated with mass media interventions. The challenges of measuring meaningful exposure are also addressed, given that the theory of change of the intervention required engagement with the plot and character identification. The evaluation employed a number of alternative options for assessing exposure and discussion of these makes a case for careful crafting of exposure measures.

**The Impact of Cell Phones on Grain Markets in Niger**

**Presenter: Jenny C. Aker**

Due partly to costly information, price dispersion across markets is common in developed and developing countries. Between 2001 and 2006, cell phone service was phased in throughout Niger, providing an alternative and cheaper search technology to grain traders and other market actors. We construct a novel theoretical model of sequential search, in which traders engage in optimal search for the maximum sales price, net transport costs. The model predicts that cell phones will increase traders’ reservation sales prices and the number of markets over which they search, leading to a reduction in price dispersion across markets. To test the predictions of the theoretical model, we use a unique market and trader dataset from Niger that combines data on prices, transport costs, rainfall and grain production with cell phone access and trader behavior. We first exploit the quasi-experimental nature of cell phone coverage to estimate the impact of the introduction of information technology on market performance. The results provide evidence that cell phones reduce grain price dispersion across markets by a minimum of 6.4 percent and reduce intra-annual price variation by 12 percent. Cell phones have a greater impact on price dispersion for market pairs that are farther away, and for those with lower road quality. This effect becomes larger as a higher percentage of markets have cell phone coverage. We provide empirical evidence in support of specific mechanisms that partially explain the impact of cell phones on market performance. Robustness checks suggest that the results are not driven by selection on unobservables, nor are they solely a result of general equilibrium effects. Calculations of the four-firm concentration index suggest that the grain market structure is competitive, so the observed reductions in price dispersion are not due to greater market collusion. The primary mechanism by which cell phones affect market-level outcomes appears to be a reduction in search costs, as grain traders operating in markets with cell phone coverage search over a greater number of markets and sell in more markets. The results suggest that cell phones improved consumer and trader welfare in Niger.
Statistical meta-evaluations of impact evaluations: a role for mixed evidences

Presenter: Tanguy Bernard

The past decade has seen a large increase in the number of statistically robust impact evaluations in the field of development. However, while important results were uncovered, it is not yet clear whether such “hard evidence” is enough to influence policy makers into ‘doing the right thing’. Even in the US where such evaluations have been implemented in Social Welfare Programs since the 1960’s, the use of the results by policy makers has had varied successes.

There are two broad reasons for that. First, aid decisions may react less to evidence than to political fashion. This may be true at government as well as development agencies levels. Second, impact evaluations sometimes lack the capacity to produce useful and usable knowledge. Indeed, while the internal validity of impact evaluation studies has greatly improved over the past few years – in particular through the use of randomized evaluations – it does not imply that external validity (the generalizability of the findings to other environments or to other program features) is satisfied. As a matter of fact, it is often argued that it is the nature of controlled experiments that makes inference “clean”, that also makes the generalization of their results to other environments problematic. And indeed, program’s effects may greatly vary across environments. Studies in the US for instance, have shown that findings from multi-sites randomized evaluation inevitably vary across sites. Similarly in Mexico, the effects of the famously evaluated PROGRESA program were three times as large in richer states than in poorer ones. If such heterogeneity can be found within countries, it is fair to suspect that results may also vary greatly from one country to another.

This paper assesses whether the usefulness of impact evaluations can be enhanced through meta-analytical approaches. We briefly review the actual debate on evaluations. For some authors, the use of randomized designs allows robust answers to questions for which previous answers rested on largely untestable assumptions. For others, the mainly statistical and microscopic nature of these studies poses the threat of “missing the forest for the trees”. We then discuss how replication of such studies could help address these criticisms – by raising the number of trees. There are however important obstacles to the promotion of these evaluations such that a large number is not likely to be available for at least a few years. We therefore discuss the extent to which internally valid studies can be combined with ‘softer’ quantitative measures in order to draw more ‘externally valid’ conclusions. A simple analytical framework is developed to show that being too restrictive may restrict ability to generalize, while being too inclusive may weaken the confidence that can be placed in the findings. However, we show that under certain conditions, the mix of evidence can lead to unbiased meta-estimates. Satisfaction of these conditions in turn requires a significant level of coordination between development agencies.
School feeding programs in Low and Middle Income Countries: evidence and implications for policy and practice

**Costs and cost-outcomes of school feeding programs in Africa**

**Presenter: Elizabeth Kristjansson**

This cost-outcome study is based on a Cochrane/Campbell review of school meals (Kristjansson, 2007). Our objectives were to: 1) provide an up-to-date, realistic estimate of the costs of school feeding 2) combine these estimates with results of the Kristjansson review to estimate all costs of school feeding programs to the World Food Program, governments, and communities and 3) provide recommendations for program implementation. Costs were calculated for four countries: Kenya, Malawi, Lesotho, and the Gambia.

School feeding programs improved attendance and math performance, produced small gains in height and weight, and show promise in improving performance on intelligence type tasks. Costs per child per year ranged from $28 to $61 dollars. The cost for an extra day of attendance per child per year ranged from $4 to $12, while the cost of an extra point on the Wide Range Achievement test ranged from $31.6 to $69.

School feeding programs can improve some aspects of the health of disadvantaged children. The cost/outcome ratios can be reduced by improving program implementation. We provide several recommendations for making school feeding programs more effective and cost effective. This is particularly important in an era where world food supplies are dwindling.

**Challenges and advantages of systematic reviews**

**Presenter: Jeff Valentine**

Systematic reviews have the potential to aid political decision making. This section of the workshop will describe the process of systematic reviewing, and show how systematic reviews are generally superior to other kinds of reviews. The Campbell Collaboration's process for helping to ensure the quality of systematic reviews will be described. Emphasis will be placed on the challenges associated with conducting systematic reviews with relevance to developing countries, including the need for more and better studies to serve as the raw materials for the reviews.

**Evaluation, propaganda, policy and practice**

**Presenter: Eamonn Noonan**

Kristjansson et al provide important evidence on the effectiveness and cost-outcomes of school feeding programs in LMIC: if well implemented, they can result in improvements in child health and development. Such programs should be retained, and imitated. How can we make this happen? This presentation examines three issues: how to publicise the evidence; how to address obstacles to further action, and how to encourage a suitable response in policy and practice.

Political stalemates and administrative barriers can block the continuation and expansion of worthwhile programs, quite apart from perennial financial constraints. But there are signs of a new pragmatism in the political sphere, and a new scepticism about simplistic, ideology-driven strategies. This may be a good time to renew our dialogue with the policy community, both North and South, to promote pragmatic solutions such as nutrition programs.

Decision-making often seems to be influenced more by studies that are highly publicised or that support prior beliefs than by systematic reviews of all available relevant evidence. In order to promote good practice we need to draw attention to the evidence base and to consider new avenues of communication.
Evidence and implementation

Presenter: Donald Bundy

A co-author of the study presented by Kristjanssen responds to the previous presentations and addresses ways to open communication channels. Solid evidence is important for policy making and program implementation; how do we get the evidence to the people who need it, and how do we ensure that it is both useful and used?

Impact Evaluation of Islamic Microfinance using Laboratory Experiments

Presenter: Mohamed El-Komi

In this paper we discuss the use of laboratory pre-tests as a mean to evaluate the impact of newly designed microfinance products that conform to Islamic finance rules. The impact of such products is evaluated with regard to their sustainability, in the absence of the use of subsidies and charity. Sustainability of microfinance projects is a main concern for practitioners and policy makers in the developmental fields. The UNCDF Bluebook (2005) asserts that more than 95% of microfinance institutions still require subsidies to cover their costs and finance their loans. Developing products that can attract commercial banks to the field has been a major challenge in the microfinance world. Therefore, these Islamic microfinance products are designed to attract commercial Islamic banks, which will have a great impact on pumping funds into this important field of development.

There are several methods to evaluate the impact of developmental programmes and products. In this paper we use an experimental economic approach to perform this evaluation. This approach is based on the use of laboratory experiments. The laboratory is an excellent setting for a testbed of new products. In the experimental lab, the environment can be controlled, and the impact of the various products compared without concerns of endogeneity or omitted variables which plague observational data. In this sense, the laboratory is like a wind-tunnel; it allows us to identify the effects of each product in isolation, controlling for other factors which might interfere, much as physicists use vacuums to examine the interaction of particles without the interference of air resistance. Of course, once these products have been tested and demonstrated successful in the lab, the next step is a field test where their performance can be evaluated “in the wild,” in the presence of endogeneity, omitted variables, and even air resistance. Nonetheless, we consider the laboratory pre-tests a critical component of evaluating the intervention.

Laboratory pre-tests are carefully designed and implemented. Participants are not paid a flat amount for their time, but instead are incentivised in a way which matches real-world outcomes. Thus, if participants default on a loan, they keep the money loaned to them (earn more) and the lender loses the money loaned to them (earn less). If the borrower is audited and caught, they pay penalties to the lender (they earn less, and the lender earns more). In this way, participants’ decisions have real (financial) consequences, just as real-world borrowers’ and lenders’ decisions have real financial consequences. The design of each of the products will be tested in a separate lab experiment. Each of these experiments will include
two treatments (control group and intervention group). In order to comply with the principles of induced valuation, participants will be paid in cash in a way which is responsive to the decisions they make in the experiment.

Our ultimate goal is not only to develop reliable and profitable Islamic microfinance products. We also wish to provide an evaluation: a side-by-side comparison between these new products and the more traditional interest-based banks, including the attractiveness of the different types of products to different populations, the repayment rates, and the risks and returns to financial organizations which offer them.

**Themes 2: Impact evaluation: Approaches and methods**

**Themes 6: Impact evaluation of policies, programs, projects, partnerships and networks**

**Session stream: Microfinance and enterprise development**

Measuring enterprise development impact in complex, resource constrained environments in South Africa – adapting the EvaluLEAD Guide from leadership development to enterprise development

**Presenter: Don Shay**

In the midst of a challenging developing country context the South African Institute for Entrepreneurship (SAIE) wants to make a difference by developing people’s entrepreneurial behaviours and strengthening enterprising competencies. Enterprise development is a complex process that defies simplistic forms of impact evaluation. SAIE has undertaken a range of impact evaluations in schools, using a combination of quasi-experimental design and qualitative approaches.

After a few years of developing and testing these impact evaluation approaches in over 75 schools, the SAIE team has concluded that we need to develop some richer, more sophisticated ways of measuring impact. A very useful evaluation tool called EvaluLEAD\(^1\) has emerged recently in the leadership development field and SAIE is adapting it for use in entrepreneurship development (\(^1\) EvaluLEAD: A Guide for shaping and evaluating leadership development programs, Grove, J.T., Kibel, B.M. & Haas, T., Sustainable Leadership Initiative, 2005.).

Enterprise development operates in a complex, open system, where many influences affect participants and human elements make predictable outcomes difficult. EvaluLEAD suggests that three levels of change should be explored:

1. **Episodic/Immediate**: more predictable, short term, cause-and-effect results following a programme’s intervention;

2. **Developmental**: more open-ended and less predictable changes occurring over time, including forward progress and setbacks, and affected by external influences and internal willingness and ability to change;

3. **Transformative**: fundamental shifts in individual, organisational, or community values and perspectives that seed the emergence of fundamental shifts in behaviour and performance; often the ’prize’ to which programmes aspire.

These three levels of change should be considered for three domains: individuals, organisations and the broader community being served. This approach leads to a three by
three programme results map. Two main forms of inquiry are considered – evidential (facts) and evocative (opinions) – for each square in the grid.

The process of working through the EntrEval framework has helped SAIE reflect on its theoretical sources and its understanding of how change happens. And a range of core indicators have been developed that serve to focus evidence collection on impact. One of the most helpful parts of the process has been the identification of people talking about their experiences and telling their stories as a particularly important way to understand SAIE’s impact.

Themes 2: Impact evaluation: Approaches and methods
Themes 6: Impact evaluation of policies, programs, projects, partnerships and networks

Session stream: Microfinance and enterprise development


Presenter: Emem Bassey Inyang

Development projects are generally greeted with great expectations but in most cases much is left to be desired. The case of ultra modern Akpan Andem market in Uyo metropolis illustrated a typical situation of abandoned adoption of infrastructural development project. A soft and hard system research methodology was adopted to resolve some of the critical indicators that can be used for learning on sustainable development and also to reduce the extent of unsustainable project. At the early phase, focus group discussion and in-depth interview with key informant provided the numerous responses that suggested reasons why the traders rejected the state government relocation policy to the new ultra modern market and its subsequent abandonment. The application of a multivariate statistic; Factor Analysis revealed that among the eighteen reasons, six significant underlying conflict resolution indicators were identified and the factor loadings of seventeen reasons were adequately distributed across the six identified underlying factors. Based on the factor loadings of each of the eighteen items across the six underlying dimensions, suggestive naming of the six factors were made and are presented as follows, factor 1: problem of misplacement of focus and precise identification/definition of the project’s priority; factor 2: influence of conflict between traditional and Christian religious beliefs; factor 3: low psychosocial adjustment and lack of participant-oriented development planning; factor 4: poor understanding and recognition of the socio-political dynamics of the social groups at the Uyo main market; factor 5: inefficiency in information educational communication from the project owners; factor 6: non-identification of the felt needs of the target population. The six underlying factors portray a lot of implication on how to minimize conflict that can be associated with developmental projects that attract the masses interest. Therefore, much conflict friendly development programming education is most desirable for development agents serving in the governmental institutions since different tier of governments will still be the drivers of development in decades to come.
Microfinance: what do we know? Learning from field experiments

*Presenters: Dean Karlan, Annie Duflo*

Microfinance has expanded at an astonishing growth rate over the last decades. In 2006 Mohammad Yunus received the Nobel price for his work, signaling international enthusiasm for microfinance as a development tool. Yet despite plethora of anecdotal evidence, we know very little about how microfinance has affected the lives of the poor, nor of the best and most beneficial ways to provide financial services to the under-banked. While the Grameen model of microfinance has proven effective at scaling up, there has been little innovation on product and program design.

Randomized Controlled Trials can help practitioners not only evaluate the impact of their programs, but make important programmatic decisions and determine what works best by conducting product innovation testing and taking controlled risks while doing so. The last few years have witnessed a large increase in the number of impact evaluations in the field. Innovations for Poverty Action has taken a lead in such research, with over 30 projects on microfinance.

This panel aims to showcase ongoing and completed research projects, and to take stock of what we have learned and what is yet to learn about microfinance operations around the world. It will draw from randomized experiments conducted by IPA in South Asia, Latin America, the Philippines, South Africa and Ghana with its academic research affiliates and research partners including centers in developing countries such as the Centre for Micro Finance in India (CMF) or ISSER (Institute of Statistical Social and Economic Research) in Ghana.

The panel will feature innovative ways to use randomized designs for improving product design and policies and will explore three broad sets of issues: what is the impact of microfinance on the beneficiaries; how to expand financial access to the poor, and how to improve the quality of financial services available to them. Key policy questions include: Do microfinance customers care about price, or are they more concerned with other features and terms of the financial service? Why are take-up rates often so low for new products? Does education and financial literacy make a difference? How does flexibility of products affect clients, and Microfinance Institutions? Is bundling credit with insurance an effective way to mitigate risk?

By comparing results from similar experiments conducted in different countries, this discussion will provide important insights for policy makers and practitioners in Africa and elsewhere.
Mexican experience on impact evaluations for 3 social programs with nutritional objectives

**Presenter: Ricardo Mujica**

The Ministry of Social Development in Mexico (SEDESOL) operates, among others, three programs with nutritional components:

Oportunidades. Conditional Cash Transfer Program with Education, Health and Food components. Aimed at poor households with over 25 million beneficiaries. Fortified Milk Program, aimed at specific members of poor households (children under 6 years of age, pregnant and lactating women, elderly) with over 2.7 million households. Food Program. In kind or cash transfer program aimed at the poorest households in Mexico that cannot comply with corresponsabilities of Oportunidades Program due to the lack of health or education infrastructure.

Since 2003, SEDESOL’S General Directorate of Evaluation and Monitoring has conducted impact evaluations for these programs. The experience accumulated over the last 5 years has recently (June, 2008) been incorporated into a book (Nutrition and Poverty: Public Policy based on evidence [http://siteresources.worldbank.org/MEXICOINSPANISH/Resources/NutricionyPobreza.pdf](http://siteresources.worldbank.org/MEXICOINSPANISH/Resources/NutricionyPobreza.pdf)) that offers a global point of view of the nutritional evolution of Mexico from 1998 to 2008 and the impact of social programs on this evolution. This book is a co-publication from SEDESOL, the World Bank and Mexico’s National Institution of Public Health.

The book’s objective is to provide evidence for public policy design on nutrition, specifically for policy aimed at poor households. It presents actual information on the nutritional situation of Mexico’s population, based on nutritional national surveys and analyses the evidence on the pertinence, targeting and actual impact of Mexico’s main social programs with nutritional objectives. The book presents evidence on important achievements on under-nourishing reduction and how these achievements are partially a direct impact of better designed social programs. Also, discusses remaining nutritional problems such as persisting anemia incidence for most social groups.

---

**Using Anthropometric Assessment of OVC to Improve Program Quality and measure program Impact**

**Presenter: Audu Muhammed Liman**

The Basic Education Support Project is an education support as well as an HIV and AIDS mitigation support project funded by PEPFAR since 2006. The project supports the Namibian
Ministry of education system through improving the capacity of the system to deal with the impact of the pandemic. One such activity is school feeding that is provided to OVC to keep them in school so they can successfully complete primary education. The project has built systems to monitor and evaluate the impact of the feeding program on the performance and health status of the OVC through an anthropometric assessment. This system is now being scaled up and turned over to them ministry. This abstract reports on the work done by the project to collect baseline data and build capacity within the Ministry of education to effectively run a school feeding program.

**Design/Methods.** In collecting anthropometric data for monitoring nutritional status a number of approaches exist. However in order to ensure consistency and wider use of data in Namibia, it was decided that the current approach used to measure under 5 children be extended to the OVC measurement. This approach was the weight for age (WfA) that is referenced in the Namibian DHS as being “One of the most commonly used reference populations and one used in this report, is the U.S. National Centre for Health and Statistics(NCHS) standard, which is recommended for use by World Health Organization (WHO).” The choice of this measurement was based on its simplicity as well as use of to chart growth of boys and girls. The project supports about 5100 OVC under 41 school feeding schemes. To avoid stigma and discrimination in the schools it was decided that all learners in the beneficiary schools will be measured and the date will be disaggregated by gender as well as OVC status. All 5100 OVC and about 4600 non OVC in the 41 primary schools were measured after the class teachers were trained and provided with the tools. They then captured the data in hard copy and kept a copy at the school, while sending the original to the project office. The data was captured into Epi Info and is now being analyzed to set baseline.

**Themes 2: Impact evaluation: Approaches and methods**

<table>
<thead>
<tr>
<th>Session stream: Participatory approaches</th>
<th>Session name: Participatory approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Session number:</strong> [31.1 ]</td>
<td><strong>Theory of Beneficiary Impact Assessment: practical insights from rural access interventions in Ratnapura, Sri Lanka</strong></td>
</tr>
</tbody>
</table>

**Presenter: Mansi Kumarasiri**

Involvement of beneficiaries in project impact evaluations, known as beneficiary/participatory impact assessment, is currently widely used in the impact studies. Ideally, beneficiary involvement gives access to information that is often difficult to be collected by an external evaluator, and provide information with greater credibility for project steering. It can also minimize the proportion of the project budget spent on monitoring so that a greater proportion could be spent on the actual intervention. The paper draws from a case study where, in theory, the circumstances were very conducive to beneficiary impact assessment, to understand what factors are crucial for this methodology to be successfully implemented.

Beneficiary impact assessment methods were used in a rural infrastructure project, implemented by Practical Action and funded by SIDA, in three rural villages in Ranthnapura, Sri Lanka. The intervention piloted a model of site selection to implement pro-poor rural access development projects and a parallel poverty impact monitoring of the intervention. Project implementation, including the monitoring and evaluation, was planned and implemented as a participatory activity. The monitoring itself was carried out at three time periods; baseline, in-process and ex-post monitoring. While the baseline and ex-post evaluations followed an external evaluator orientation, the in-process monitoring component piloted beneficiary impact monitoring/assessment complementing the community participation methods used in the actual intervention to construct rural roads. Groups of three to four
selected participants were appointed in each of the three villages and trained on the tools to carry out the monitoring. Within each group, one member was responsible for reporting on a given tool on a regular basis. Because the project involved a priority need for the community and also because project implementation was participatory, the circumstances were particularly conducive to the use of beneficiary impact assessment methodologies.

The paper analyses the experiences of this case study to draw out insights for practitioners of beneficiary impact assessments, particularly in cases of participatory infrastructure development projects, to improve the likelihood of successful implementation of the methodology. Although the methodological objectives relating to low cost monitoring were easily achieved, the objectives obtaining adequate and timely information for project steering and to understanding the overall impact of the intervention, depended mostly on, the environment within which the methodology was implemented such as the features of each intervention site, the carrying capacity of a group to implement both a participatory construction as well as the monitoring activity, and the composition of the beneficiary group responsible for providing leadership to the monitoring exercise.

### Themes 2: Impact evaluation: Approaches and methods

**Session number: [31.1]**

**Session stream: Participatory approaches**

**Session name: Participatory approaches**

### Developing a participatory impact assessment approach and action research culture within a communication for social change organisation in Nepal

**Presenter: June Lennie**

Developing an effective impact evaluation framework, managing and conducting rigorous impact evaluations, and developing a strong research and evaluation culture within development communication organisations presents many challenges. This is especially so when both the community and organisational context is continually changing and the outcomes of programs are complex and difficult to clearly identify.

This paper presents a case study from a research project being conducted from 2007-2010 that aims to address these challenges and issues, entitled Assessing Communication for Social Change: A New Agenda in Impact Assessment. Building on previous development communication projects which used ethnographic action research, this project is developing, trailing and rigorously evaluating a participatory impact assessment methodology for assessing the social change impacts of community radio programs in Nepal. This project is a collaboration between Equal Access – Nepal (EAN), Equal Access – International, local stakeholders and listeners, a network of trained community researchers, and a research team from two Australian universities.

A key element of the project is the establishment of an organisational culture within EAN that values and supports the impact assessment process being developed, which is based on continuous action learning and improvement. The paper describes the situation related to monitoring and evaluation (M&E) and impact assessment before the project began, in which EAN was often reliant on time-bound studies and ‘success stories’ derived from listener letters and feedback. We then outline the various strategies used in an effort to develop stronger and more effective impact assessment and M&E systems, and the gradual changes that have occurred to date. These changes include a greater understanding of the value of adopting a participatory, holistic, evidence-based approach to impact assessment.

We also critically review the many challenges experienced in this process, including: (i) tension between the pressure from donors to ‘prove’ impacts and the adoption of a bottom-up,
participatory approach based on ‘improving’ programs in ways that meet community needs and aspirations; (ii) resistance from the content teams to changing their existing M&E practices and to the perceived complexity of the approach; (iii) lack of meaningful connection between the M&E and content teams; (iv) human resource problems and lack of capacity in analysing qualitative data and reporting results; (v) the contextual challenges, including extreme poverty, wide cultural and linguistic diversity, poor transport and communications infrastructure, and political instability; and (vi) a general lack of acceptance of the importance of evaluation within Nepal due to accepting everything as fate or ‘natural’ rather than requiring investigation into a problem.

### Themes 2: Impact evaluation: Approaches and methods

**Session stream: Participatory approaches**

**Session name: Participatory approaches**

**Session number: [31.1]**

**Multi Level Participatory Midterm Review: An Innovative Approach in Evaluating Project Performance, Outcome and Impacts**

**Presenter: Orlen Ocleasa**

This paper I will be presenting shall share the experiences from developing countries in Asia Pacific region such as the Philippines on evaluation of development projects at the middle of the implementation of the project life as well as the attempt to innovate a new approach in midterm evaluation. It shall reveal consolidated & comparative experiences and lessons learned by various external M & E consultants/ experts in the midterm evaluation of projects in the Philippines and neighboring Asian countries funded by IFAD (International Fund for Agricultural Development), USAID, USDA (US Department of Agriculture), GTZ (German Cooperation Agency), World Bank, Asian Development Bank, AusAID and other donors. Consultants who were hired as external experts by these donor or funding institutions where interviewed regarding their experiences, learnings and recommendations when they conducted Midterm Evaluations. Their answers were consolidated together with some data from various literature and presented as Part I of this paper which shall be sub-titled I. Midterm Evaluation: Experiences from the Philippines and Asia Pacific regions

The Part II of the paper sub-titled as II. Participatory Impact Evaluation shall discuss the practices of PIE in the Philippines and neighboring countries. It shall also share the experiences of the presenter in conducting Participatory Impact Evaluation in his previous project funded by IFAD, the Western Mindanao Community Initiatives project (WMCIP) which was implemented in conflict affected areas.

Part III shall present an experience of trying develop a new midterm evaluation methodology to find a more inclusive and meaningful process to assess the project performance, outcomes and impacts (even the initial or emerging impacts) at or just beyond the mid-point of the program implementation. This is the Multi Level Participatory Mid-Term Evaluation, an innovation from the traditional midterm review using 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 if it is still too early. 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 fora. All their outputs were consolidated into one report which is now the Participatory Midterm Evaluation Review report.
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 office in Cebu City, Philippines but project was implemented in a nationwide scale.

**Themes 2: Impact evaluation: Approaches and methods**

**Session stream: Participatory approaches**

**Session number: [31.1]**

**Session name: Participatory approaches**

**Comparative Constituency Feedback – a practical way of facilitating whole system learning and mutually accountable relationships for outcomes**

**Presenter: Andre Proctor**

Comparative Constituency Feedback (CCF) surveys and reports provide organizations with transformative information about how their policies and practices are perceived and experienced by their primary constituents, and how this perception compares with that of other similar organizations in their field.

CCF surveys are not just a one time, one way data extraction exercise. They create a space in which the least powerful constituents of a social change process can safely and anonymously share their honest perceptions of the practices, systems and behaviours of more powerful constituents like funders or implementing NGOs. They place hard qualitative and quantitative data on the table around which grantmakers, intermediaries and recipients can engage in ongoing purposeful learning dialogue to enhance their relationships and their contributions to meaningful social change.

When a powerful actor (like a funder or an NGO working with poor or vulnerable people) commits to listening and responding to this kind of feedback in an open and transparent way, the power imbalance begins to shift subtly creating conditions for real generative dialogue. New kinds of purposeful relationships become possible.

**Themes 2: Impact evaluation: Approaches and methods**

**Session stream: Participatory approaches**

**Session number: [31.2]**

**Session name: Participatory approaches**

**Constituency Voice - a key element in impact evaluation**

**Presenters: David Bonbright, Lawrence Haddad, Anabel Cruz, Jacinto C. Gavino, Catherine A. Odaro-Hoppers, Alnoor Ebrahim**

This session will explore how rigorous approaches to constituency feedback can enhance the way we understand and realize impacts. Taking constituency feedback seriously at every stage of the operational cycle -- planning, implementation, assessment and reporting -- enhances progress on intended outcomes. The presentation will unpack 'Constituency Voice' methodology, as it is called, with a focus on its implications for impact evaluation. Examples from recent investments from Bill & Melinda Gates Foundation, WK Kellogg Foundation, and others social investors will illustrate how the context for impact evaluation is shifting to one in which improvement and learning are prioritized. The talk will also interpret the significance of 'customer satisfaction' in the business world for impact evaluation, building on the article that David Bonbright is currently writing with Jamey Power of JD Power and Associates. The roundtable will count with the participation of renowned experts in the field of social change and innovation from the four corners of the world.
Participatory Monitoring and Evaluation for Poverty Reduction

Presenter: Sindu Workneh

There is a broad consensus that achieving sustained development in developing countries depends on breaking the vicious circle of poverty and low level growth. To break this vicious circle, different development policies and strategies are being implemented by development practitioners. However, the level and intensity of poverty in Africa shows an increasing trend resulting in more than 1.2 billion people living on less than $1 a day. Moreover, poverty is much more than the mere statistics. It includes, amongst many things, lack of hope; people’s inability to achieve their full potential; lack of full participation of local people in development activities; absence/low level of empowerment of local people; and missed opportunities (Mule, 2006).

Monitoring and Evaluation (M&E) of development activities is a key element in poverty reduction strategy of developing countries as it provides the concerned bodies with better means for learning from past experiences, improving service delivery, planning and allocating resources, and demonstrating results as part of accountability to key stakeholders (World Bank, 2004). While M&E can be done by ad-hoc “expert” approach, Participatory Monitoring and Evaluation, hereinafter PM&E, offers development activities a host of opportunities for improving the performance of poverty alleviation programs. PM&E is people centered and entails the active involvement of project stakeholders and beneficiaries to generate a sense of ownership in the M&E results and recommendations. PM&E approaches encompass a wide and expanding range of philosophies, tools and methodologies that help in strengthening poverty alleviating programmes to achieve impact. The PM&E help in improving the performance of development and poverty alleviation programs by enhancing local learning, management capacity and skills, build partnerships and sense of local ownership over projects, and ultimately empower local people to make their own decisions about the future (Sartorius, 2003).

Given the advantages of PM&E, there is no doubt that it should be an integral part of project planning and implementation. Designing PM&E systems in development projects at the inception stage will increase the likelihood that PM&E is not an after-thought, rather is fully integrated in project plans and implementations. In the process, important PM&E benefits such as participatory learning and action aimed at project improvement are realized.

This paper addresses the rationale of integrating PM&E in development projects and programmes, assesses the different methods and practices of PM&E using case studies and concludes that there is no single recipe to integrating and implementing PM&E. Rather, PM&E is context specific, rooted in the concerns, interests and problems of program end-users and in the process impact the community at large in reducing poverty. In addition, undertaking PM&E does not mean involving everybody in all decisions at all times. Rather, it means making an optimal use of the identified stakeholders by ensuring that different interests of stakeholders are met in different phases and forums of the PM&E process.
Pró-Conselho Indicators Tool. A participatory methodology for planning and evaluate results

Presenter: Ana Cristina Valente Borges

The role of the third sector organizations, understood here by formal institutions and private, distributors of non-profit, autonomous, voluntary and public purpose (7 & SALOMON ANHEIER, 1997), whose programs aim to meet basic social rights, fight against exclusion and protect the ecological heritage (SZAZI, 2001), is a phenomenon that arouses the interest of researchers from different sectors.

One of the prisms under which this universe can be seen is the relationship (access and use) with information and activities of monitoring and evaluation established in that segment, in particular processes aimed at the continuous monitoring of results and use of methodologies based in techniques for monitoring and evaluation as a management tool, dissemination of information and the construction of applied knowledge.

This paper aims to present a tool designed both to contribute with development of strategic planning and management skills and to facilitate the monitoring and evaluation of the impact of a program implemented in Minas Gerais, Brazil, in favor of strengthening the system of guaranteeing the rights of children and adolescents: the Pró-Conselho Indicators. In addition to measuring the progress achieved by the program as a whole, this approach also enables councillors of Program beneficiary Councils to work at each of the dimensions that seems to be fundamental to the management of this type of institution, considering its planning for the short, medium and long term. The baseline that references this work is that advocated by evaluation and monitoring theories, which deals with the application of formative assumptions and techniques considering the strategic planning of activities and control of projects. The goal is the understanding of the role that the relationship information - monitoring - evaluation - the organization can play in planning, management and strengthening of this organizations which have the common good as an ultimately aim. The theoretical foundation will be a conductor of the description and analysis of the problem centered on the relationship between information (access and use) by the Councils of the Rights of the Child and Adolescent, in Brazil. The councils are a public policy formal institution expected to be organized in all the 5.483 municipalities existing in Brazil with representatives from the public sector and from the civil society. The program as a whole is structured in a inter-sector and inter-organizational partnership base, involving multiple actors. The Pró-Conselho indicators includes over 14 dimensions of activities and routines usual for those institutions and considers four types of indicators: development phases of management of councils and public policies implementations (indicators of deepness); quantitative indicators; binary indicators (yes or no); and scale indicators with multiple alternatives.
**Using participatory methods to assess impacts**

**Presenter: Robert Chambers**

The opportunities presented by the explosion of participatory methodologies (PMs) are still largely overlooked. The dimensions are paradigmatic, professional, institutional and personal. PMs shift the balance from the rigours of linear reductionism to a new rigour of synergies with complexity science. Evidence and argument will support five propositions:

- Learning and changing are significant impacts
- Questions of power and of whose realities count – those of professionals or those of local people - are fundamental
- Numbers generated through PMs can fit and express local complexity and diversity, quantify the qualitative and be commensurable for statistical analysis
- PMs for impact assessment can be win-wins, giving insights inaccessible with traditional methods, and enhancing impact by informing and empowering local people.
- Obstacles to PMs include teaching and training, mindsets and reflexes, misapplied scientific methodology, and the power of a conservative establishment

**Market Access for Africa's Transformation and Development: A perspective for evaluating optimal impact on development**

**Presenter: Hilary Nwokeabia**

This paper provides a perspective to impact evaluation and policy re-orientation for some of the puzzling dynamics of market access initiatives for Africa's export competitiveness, in global trade. At the beginning of the 1980s, African countries contributed marginally and narrowly to the global trade. As the difficulties with trade got deep, some partners of the region, particularly the Triad (the United States, Japan and European Union) adopted legislated trade advantages, granting up to 30% cost advantage and 6,400 items duty-free and quota free, in some cases, to regional exporters against exclusion from, and narrow participation in the global economy. Over time, preference erosion has meant that the impacts of some of the initiatives, particularly those related to extensive margins of trade, were short-lived as their provisions ran into other problems or expired. The paper applies the logic of Area of influence (AOI) and Influence mobility (IM) models to add a framework for evaluating "magnification" effects to the complimentary/compensatory relationship between the legislated export advantages and factor input export competitiveness, and the erosion of these advantages. In the algorithm of AOI and IM models, the binding theme is that for trade initiatives to be sound and economic growth sustainable, they must not only meet short term intensive trade
margins, but also help catalyze long term factor input competitiveness and extensive export margins. Events over the past years strongly suggest that the recent export boom due to GSP initiatives proved to be short-lived because legislated advantages were dominating in their effects on the region's extensive margin of exports related to these initiatives. The initiatives have not been complimented by domestic base knowledge and an iterative productive development process through vertical specialization. Thus, there are many steps African Governments must take to boost their global trade competitiveness and contributions.

**Themes 2: Impact evaluation: Approaches and methods**

**Themes 6: Impact evaluation of policies, programs, projects, partnerships and networks**

**Session stream: Private sector development**

---

**Do Public Works Crowd Out Private Transactions: Evidence on Labor Market Participation, Credit and Transfers in Ethiopia**

**Presenter: Daniel O Gilligan**

**Co-authors: John Hoddinott, Neha Rati Kumar, Alemayehu Seyoum Taffesse**

Social protection programs are often justified as a necessary public response to market failures. Insurance and credit markets are not sufficiently developed to insure poor households against risk or to support investment. Program transfers partially substitute for these missing private markets. However, these transfers may also replace private transactions, particularly credit transactions or transfers between households. If this crowding out effect is large, it may weaken credit markets or social networks.

Increasingly, transfers from social protection programs are conditional on an activity of the household, such as sending a child to school or visiting health clinics. Public works programs offer a longstanding form of conditioning, providing transfers in exchange for labor. If not carefully designed, public works may crowd out labor supplied in the private labor market. This leads to inefficiency unless the returns to the public goods (e.g., infrastructure) provided through public works are high, or in the atypical case in which public enterprises are more efficient than private ones.

This paper examines the empirical evidence on whether private transactions in labor and credit markets or transfers between households are crowded out by public works conducted under Ethiopia’s Productive Safety Net Program (PSNP). The PSNP is one of the world’s largest public works projects, employing more than 5 million people annually. By design, the program demands substantial labor inputs from beneficiary households, with a quota of 36 labor days per household member over a six month period each year.

Using a household panel data set on 3700 households collected in 2006 and 2008 across four regions in Ethiopia for the purpose of evaluating the public works program, we measure the program's potential role in crowding out labor market participation. We use difference-in-differences matching estimators to study the effect of the PSNP on the change in probability that working-age members of beneficiary households supply labor to the private labor market over this period. Results based on propensity score matching and nearest neighbor covariate matching are compared and we discuss the strengths of these estimators in drawing conclusions from the estimates. We look for heterogeneity of treatment effects by gender, by household labor endowment, and by poverty status. By comparing households with low and
high rates of participation in the program, we examine endogenous dose effects of the labor quotas and estimate the labor requirement at which crowding out begins. We attempt to identify the exogenous component of this effect by instrumenting for labor hours supplied to public works using household demographic characteristics, distance to markets, and proxies for local labor demand.

We also examine the impact of the PSNP on use of credit and on net receipt of private transfers. Isolating the crowding out effects on credit use is complicated because the effect of program transfers on credit demand is ambiguous. We use the panel data on nonbeneficiaries to predict credit market entry by PNSP beneficiaries in the absence of the program. We compare these predictions to beneficiaries’ observed credit market behavior to estimate the program’s effect on crowding out of private borrowing.

Impact evaluation Approaches to Non-Traditional Development Assistance: The case of China’s mining investment in the Democratic Republic of Congo (DRC)

**Presenter: Immaculate Dadiso Motsi**

Official Development Assistance (ODA) is defined as government aid (made up of grants or concessional loans) to developing countries designed to promote the economic development and welfare of recipient countries. Along with the traditional sources of ODA such as that from the OECD, there has been a recent increase in what has been described as non-traditional development assistance. Coming from countries such as Saudi Arabia, Iran, South Korea, India and China, this non-traditional development assistance emphasizes south-south cooperation, and seeks to promote mutual benefits.

With this in mind, this paper presents the key difference between traditional and non-traditional sources of development assistance highlighting the implications of these differences on impact evaluation. Here, it is shown that the pursuit of mutual benefits in non-traditional development assistance calls for the provision of equal weight in the impact evaluation process to not only the benefits, outcomes and effects of development assistance but also to the potential costs of such transactions on the aid recipient country. In essence, what non-traditional development assistance calls for is an evaluation of the net social benefits vis-à-vis the net social costs to the country receiving development assistance. Methodologically, this would entail not just establishing causality and considering counterfactual scenarios, but the inclusion the active pursuit of identifying negative externalities. The rationale of this approach stems from the fact that the country in question will not simply be a passive receptacle of development funds as it is in the case of traditional ODA but rather an actively engaging partner from whom equal benefit is expected by the donor.

Nowhere does this scenario more fully manifest itself than in the case of China’s development assistance to the Democratic Republic of Congo (DRC). Meeting all the defining features of non-traditional development assistance, the case study of China in the DRC gives some telling insights into both the role of impact evaluation in improving development effectiveness and impact evaluation.
Realities of Rehabilitation and Resettlement Policies and its Implementation Strategies: An Assessment of a Super Thermal Power Project in India

Presenter: Uma Chatterjee

Most large-scale development projects in India have been planned and executed through a centralized, hierarchical and capital-intensive approach, which has assumed that technology has the answer to all development problems. Centralized command and control system has perhaps been successful for achieving project objectives in terms of physical commodities like generation of thermal power, hydroelectric power, irrigation canals etc. However, the approach has almost completely failed in resettling and rehabilitating the displaced population who have suffered intolerable injustice and trauma.

The development projects have already displaced 21,300,000 people from 1951 to 1990. At present, the overall number of displaced exceeds 30,000,000. More than 150,000 people have been displaced in the last 30 years to make way for power generation projects till date. Super Thermal Power Company (STPC)(fictitious) is the biggest thermal Power Company of India claims that they have developed a policy to improve or at least regain the previous standards of living earning capacity and production levels of the displaced people through a participatory process.

The paper provides a basis for determining whether the Project Affected People (PAP) due to the setting up of the project have been actually provided with the means to restore their former living standards, thereby providing an understanding whether the vulnerability has increased or decreased and whether the degree of sustainability has improved or not. It also helps to assess the changes and outcomes that have occurred in the social fabric of the affected community.

It is an endeavor to shed some light on the extent of success of the STPC officials in implementing R & R. Despite all their sincere efforts in development activities, the authorities have to face public wrath and retaliation. Why do such problems emerge? Have the authorities tried to listen to the cries of the PAPs and PAFs and to understand sympathetically? What lie underneath? Have the concerned officials entered through the peoples’ door to solve the problems or have they gone through the STPC door for their own protection? This paper highlights some of these questions and possibly also offers some answers.
Global Partnership for Development: Implications for the evaluation of development and development aid

Presenter: Sheelagh O’Reilly

In recent years there are increasing arguments that there is an emerging body of international law that ‘all states are bound by legal obligations to cooperate in ensuring the right to development and the universal realization of basic socio-economic rights’. It has been recently argued that the international development community have agreed a contract and that this ‘contract is more aid and country ownership in exchange for a renewed focus on measuring, disseminating, and acting on results.’ Questions of ‘contracts’ provoke concern over the who the parties to the contract are, what the roles and responsibilities the parties and how will the contract be ‘enforced’ if one party does not comply with the terms and conditions. The language of ‘contract is not used in the Monterrey Consensus where there is a clearly articulated need for partnerships to increase the impact of development programmes including those funded by ODA i.e. ‘development aid effectiveness’. The Monterrey Consensus also recognizes the need for a ‘holistic approach to the interconnected national, international and systemic challenges of financing for development’ which it also recognized includes the promotion of ‘national and global economic systems based on principles of justice, equity …. i.e. development effectiveness.’ In the International Law of Development the focus is not on contracts per se but on the balance between state sovereignty and the interdependence of states around common interests and values articulated in the UN Charter and subsequent treaties, conventions and processes developing international law.

This paper will therefore examine the implications of the MDG8 focus to ‘Develop a Global Partnership for Development’ and how the structural aspects of international cooperation under International Development Law such as ‘mutual accountability’, ‘maximum availability resources’, and ‘human rights law’ have for the development of monitoring and accountability procedures implemented by states in partnership with citizens and other national and international organizations. These approaches will examine the different issues raised by ‘development aid effectiveness’ and the much broader ‘development effectiveness’.

Impact evaluation: Towards a Rights based Approach

Presenter: Milindo Chakrabarti

India is a signatory to the “Declaration on the Right to Development” adopted by General Assembly resolution 41/128 of 4 December 1986. As per the website of the United Nations Commissioner for Human Rights located at Geneva [available at http://www.unhchr.ch/development/approaches-04.html as on 5th December 2008] rights-
based approach to development is a conceptual framework for the process of human development that is normatively based on international human rights standards and operationally directed to promoting and protecting human rights.

Essentially, a rights-based approach integrates the norms, standards and principles of the international human rights system into the plans, policies and processes of development. The norms and standards are those contained in the wealth of international treaties and declarations. The principles include equality and equity, accountability, empowerment and participation. A rights-based approach to development includes the following elements:

- express linkage to rights
- accountability
- empowerment
- participation
- inclusion of the excluded (equity)
- transparency

Impact evaluation, on the other hand, ‘assesses changes in the well-being of individuals, households, communities or firms that can be attributed to a particular project, program or policy’ [available at http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTISPMA/0,,menuPK:384339~pagePK:162100~piPK:159310~theSitePK:384329,00.html#whatis as on 5th December, 2008]. It thus captures the spirit of accountability in development intervention in quite an objective fashion and can contribute meaningfully to the understanding if a particular programme, project or policy has really enhanced the right to development of the vulnerable groups in India.

Given the Millennium Development Goals that attempt to reduce the incidence of poverty in a targeted manner – an effort close to the spirit of rights based approach to development -- and recent institutional changes in India to acknowledge right to education and employment on the one hand and right to information on the other, challenges to professionals engaged in impact evaluation have increased manifold. The proposed panel would like to examine the linkage between impact evaluation and rights-based approach to development and identify the processes that can contribute meaningfully towards strengthening the bond from an Indian perspective.

The key issues to be discussed are:

- What new conceptual frameworks or models need to be developed and used to capture direct and indirect impact of development attempted by state, market and other agencies concerned on the life of disadvantaged and vulnerable sections of society deprived of basic human rights and entitlements?
- How can impact evaluation of policies, programmes or projects contribute meaningfully to the understanding of a development process undertaken in India?
- Does it help understand if the spirit of ‘development as right’ is being advanced or not in the poverty reduction strategies identified?
- What are the methodological challenges involved in identifying their impact on the ‘bundle of rights’ available to an Indian citizen – particularly the vulnerable ones?
- What are the expectations of the practitioners of “rights based approach to development” from the experts engaged in impact evaluation?
- How to hold responsible authorities accountable and punish for nonfeasance, misfeasance and malfeasance at various levels that amounts to violation of human rights?

The proposed panel will involve participation by
1. Prof. Arjun Sengupta, Member of Parliament and Chairman, Center for Development and Human Rights.
2. Dr. S.P. Pal (Former Advisor, Evaluation, Planning Commission, India),
3. Prof. Samar K. Datta, Chairman, Centre for Management in Agriculture, Indian Institute of Management, Ahmedabad.
5. Dr. Milindo Chakrabarti, Director, Development Evaluation Society of India (DESI)

| Themes 1: Improving development effectiveness: the role of impact evaluation |
|-----------------------------|-----------------------------|
| Session number: [34]        |

| Themes 2: Impact evaluation: Approaches and methods |
|-----------------------------|-----------------------------|
| Session stream: Right based approaches |
| Session name: Right based approaches |

Panel: A South Asian Perspective, “Institutionalizing Impact Evaluation: Demand, Production and Use”?

Presenters: S.P. Pal, Velayuthan Sivagnanasothy, Ghulam Mustafa, Karma Galleg, Servash Kumar, Milindo Chakrabarti

Persistent poverty is a real problem plaguing the economies of South Asia. The Millennium Development Goals Report 2007, published by UNDP reports that

- 29.5% of the population of this region was suffering from extreme poverty in 2004, next only to sub-Saharan Africa in terms of regional concentration. Incidence of extreme poverty however, has been showing a decreasing trend over time.
- A whopping 46% of the children under five in this region are underweight, -- a scenario more dreadful than that prevailing even in sub-Saharan Africa.

Investments in development intervention are also not very insignificant in this region. A multi-pronged strategy involving government, multilateral agencies, bilateral donor organizations and global civil society organizations are engaged in such investments. Impact evaluation of such investments can play a very vital role in generating innovative models of interventions for the future drawing upon the lessons learnt from the past exercises.

The role of evaluation is well recognized in the planning process of South Asian countries. Every country in this region is characterized by the existence of specialized agencies to evaluate “development interventions”. While India instituted “Programme Evaluation Organization” within the Planning Commission as the apex evaluation institution in the country, Sri Lanka has entrusted such task with the Department of Foreign Aid and Budget Monitoring working under the Ministry of Plan Implementation. Evaluation exercises in Pakistan are mainly carried out by the Planning Commission and Pakistan Institute of Development Economics. In addition evaluation activities are carried out at several structural levels – ministry, province, district etc. Research Organizations and donor agencies also participate in evaluation exercises as per their respective mandates.

The proposed panel will involve participation of practicing and retired evaluators from India, Bhutan, Sri Lanka, Bangladesh and Pakistan to analyse the status of impact evaluation in their respective countries. The key issues to be covered would relate— but not necessarily confined— to the following concerns prevailing in the South Asian countries vis-à-vis impact evaluation:

- What is the status of impact evaluation today in the concerned countries?
• Who are participating in impact evaluation exercises?
• Who are demanding impact evaluation studies?
• What are the existing institutional strengths, both domestic and in terms of support from international organizations?
• What are the institutional bottlenecks to be removed to make impact evaluation a more effective tool for assisting future investment decisions? What are the challenges faced in Institutionalisation of IE in the government?
• What are the capacity development initiatives in the South Asian countries vis-à-vis impact evaluation and what are the future plans?
• How would the Impact Evaluation being used or planned to be used in the planning, budgeting and policy making process (i.e. the feedback arrangements)
• How can IE be a tool to improve development effectiveness?
• What are the methodological concerns of IE (e.g. Randomised Control Trials, Multiple methods) and factors that influence such concerns (e.g. Cost effectiveness)?

Dr. S.P. Pal (Retired Advisor, Evaluation, Planning Commission, India), Mr. Velayuthan Sivagnanasothy, Director General, Department of Foreign Aid and Budget Monitoring, Ministry of Plan Implementation, Colombo, Sri Lanka, Mr. Ghulam Mustafa, Earthquake Reconstruction and Rehabilitation Authority, Prime Minister’s Secretariat (Public), Islamabad, Pakistan, Mr. Karma Galleg, Government of Bhutan, Advocate Afsana Wahab, Executive Director, Center for Woman and Child Development, Dhaka, Bangladesh, Mr. Servash Kumar, Deputy Director General, National Sample Survey Organization (formerly associated with Programme Evaluation Organization, Planning Commission, India) have agreed to be on the panel. Dr. Milindo Chakrabarti, Director, Development Evaluation Society of India (DESI) will coordinate the panel.

Themes 2: Impact evaluation: Approaches and methods
Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks

Session number: [35.1]
Session stream: Rural development
Session name: Rural development

Impact evaluations of policies, programs, projects, partnerships and networks
Presenter: Martin Ravallion
Co-authors: Shaohua Chen, Ren Mu

The paper re-visits the site of a large, World Bank-financed, rural development program in China, 10 years after it began and four years after disbursements ended. The program emphasized community participation in multi-sectoral interventions (including farming, animal husbandry, infrastructure and social services). Data were collected on 2,000 households in project and non-project areas, spanning 10 years. A double-difference estimator of the program’s impact (on top of pre-existing governmental programs) reveals sizeable short-term income gains that were mostly saved. Only small and statistically insignificant gains to mean consumption emerged in the longer-term-though in rough accord with the average gain to permanent income. The use of community-based beneficiary selection greatly reduced the overall impact, given that the educated poor were under-covered. The main results are robust to corrections for various sources of selection bias, including village targeting and interference due to spillover effects generated by the response of local governments to the external aid.
Approaches to evaluating the impact sustainability of completed rural development projects

**Presenter: Yuriy V. Nesterov**

The proposed paper will examine approaches and challenges to evaluation of sustainability impact of rural development initiatives by example of Heifer International projects in Eastern Europe. Over a long period of operations Heifer International has accumulated a number of "historic" projects, which are projects that ended long time ago. Because of the nature of the Heifer projects, particularly because of the Passing Of the Gift (POG) mechanism, the impact of Heifer activities are visible in a long-term period. That is why there is a need to develop and test a mechanism for evaluation of long term impact of Heifer projects. The new evaluation model may be based on the evaluation approach developed by Evaluation Center of Western Michigan University for active and recently completed projects. After testing on two or three projects and further improvement, the model may be used to evaluate impact of majority of historic projects. The findings will be used to improve design of new projects and promote Heifer’s mission and for replication by other organizations that work in field of rural development.

Building the Sustainable Rural Development Index in Asia and the Pacific: Scoreboard on the Levels of Development of Rural Communities and Country Performance in Rural Development

**Presenter: Eva Benita A. Tuzon**

The rise on the use of indexes is noticeable with the increasing ease in information technology and communication. Likewise, the indexes come like capsules to remedy or invigorate our need for synthesized accounting on the various directions, achievements including slippages, all glued in the philosophy of stewarding universal goods and values. This had been manifested in the strong demand of different development stakeholders: government, civil society and business sector, including global donor communities, for results-based management critical part of which involves impact evaluations.

The paper takes off from the desire to establish an index that provides knowledge on the different levels of development of many rural communities, which development agencies claimed they have reached out. An index that would ultimately rank the different countries that have been implementing rural development programs/projects in the region. The envisioned Regional Sustainable Rural Development Index [RSRDI] is a self-conscious report...
card that aims to stand face to face with the existing indexes that measures competitiveness in terms of economic performance; performance in terms of how much was contributed to the basket of global donor funds, and other critical information that pertains to development. However, its rationale is simply to be accountable on state policies that are implemented and presumed significant to the overall framework of a sustainable rural development.

---

**Themes 2: Impact evaluation: Approaches and methods**

<table>
<thead>
<tr>
<th>Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks</th>
</tr>
</thead>
</table>

**Session stream:** Rural development

**Session number:** [35.2 ]

**Session name:** Rural development: African experiences

**Impact evaluation of the financial and educational project in the rural province of Ganzourgou**

**Presenter:** Daniel Thieba


Les questions évaluatives au centre de l’étude d’impact sont de deux ordres : (i) dans quelle mesure le Projet a pu permettre la réduction des conflits fonciers et/ou à réduire le temps et le coût pour la gestion des conflits, (ii) dans quelle mesure le Projet a pu permettre un essor des investissements en milieu rural Les autres aspects de l’évaluation concernent notamment l’usage des acquis du Projet dans une Politique Nationale de sécurisation foncière.

Pour mener à bien cette évaluation, l’auteur du présent article a mis en place un dispositif d’enquêtes spécifiques:

- collecte des données sur les conflits fonciers dans la zone d’intervention et leur comparaison avec une province voisine non touchée par l’intervention
- collecte d’enquêtes sur le rythme et l’évolution du volume d’investissements en milieu par les exploitants agricoles dans la zone d’intervention et leur comparaison avec une province voisine non touchée par l’intervention

L’objet de la communication se situe aux niveaux suivants:

- présenter la méthodologie mise en œuvre ; dans quelle mesure obéit –t-elle aux critères de rigueur dans la conduite d’études d’impact ? Comment peut-on conduire une évaluation d’impact rigoureuse sous contraintes (budget, délai, absence de certaines données de référence)
- présenter les résultats et répondre aux questions suivantes: quelle est la valeur ajoutée de cette évaluation rétrospective par rapport aux évaluations qui ont déjà eu lieu (évaluation à mi parcours et évaluation finale) ; quelle est l’utilité des résultats de cette évaluation ? Quelle usage le client (l’AFD) et l’Etat envisagent –ils faire des résultats?
Impact Evaluation of the PNIR programmes in Senegal

Presenter: Samba Mbaye

We measure the impact of the PNIR programme on abandon rate in primary and secondary schools in rural Senegal using a Hausman-Taylor IV method. Results show that the programme has a positive but non statistically significant effect on this abandon rate. Control variables such as the age of household’s chief or the age of children, as well as the number of children in primary school are mostly significant. We observe also that the desertion rate is statistically as higher for girls and boys at the primary level. At the secondary level however, this rate is slightly higher for boys.

This paper is organized as follow. The first section displays the conceptual framework of impact analysis, the second, the choice of the counterfactual. The third part of the article presents descriptive statistics and the data. The last section discusses the results and concluding remarks.

Evaluating empowerment in the rural small-holder extension context: Lessons from the field

Presenter: Deborah Duveskog

There is growing recognition globally for the importance of community empowerment, particularly in the context of demand-driven extension services. However, a very limited number of rigorous evaluation studies exists that document extension and development outcomes and impacts in terms of community empowerment. Most evaluation impact studies related to extension focus on technological change. This study helps to fill this gap by providing practical experiences in efforts to evaluate empowerment with a mixture of qualitative and quantitative measures that are both rigorous and broad in scope and provides a conceptual framework for measurement of community empowerment.

Specific objectives were to; (1) define concrete indicators of empowerment, based on the perception of small-holder farmers themselves, (2) develop, test and assess both qualitative and quantitative processes and tools for the evaluation of empowerment.

Methods used for the study include a combination of quantitative research tools including a survey and qualitative data through focus groups with farmers and local extension staffs. The empirical evidence base for the study is the East African Farmer Field School programme in Kenya, Uganda.
and Tanzania. Direct evaluation of empowerment was made by assessing: (1) Whether an opportunity to make a choice exists, (2) Whether a person actually uses the opportunity to choose (use of choice), (3) whether the choice resulted in the desired result (achievement of choice).

Through a range of discussion and reflection events at both individual and collective level were defined based on farmers and extension actors own definitions and perceptions of empowerment. Thereafter a survey questionnaire was developed for data collection on the defined indicators. Aspects related to personal perceptions and attitude were framed in the format of Likert items where respondents specified their level of agreement or disagreement to particular statements. The survey was implemented among graduated FFS members in the three countries. Methods of analysis included descriptive statistics and the use of principal component analysis to cluster questionnaire items into empowerment factors. For purpose of triangulation and providing a richer picture of the empowerment process, in-depth discussions and meeting were held with farmer groups and extension facilitators to reflect on personal expressions of the defined empowerment indicators.

This paper provides concrete experiences and lessons learned from the evaluation process undertaken. Specific indicators defined for evaluation of empowerment include; farmers ability to articulate choice and demand services, ability to negotiate with traders and external service providers, household gender roles, community solidarity and collective action. This information will thus be valuable for the research community as well as extension and development practitioners engaged in the evaluation and analysis of projects and policies aimed at community empowerment.

Themes 2: Impact evaluation: Approaches and methods
Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks
Session number: [35.3]
Session stream: Rural development
Session name: Rural development

An application of the quasi-experimental design of Impact Assessment in Sindh Rural Support Programme, Sindh Pakistan

Presenter: Fazal Ali Saadi

The Rural Support Programmes Network (RSPN) has recently conducted an impact assessment of one of its member organisation Sindh Rural Support Organisation (SRSO).

Impact Assessment of SRSO programs aims at gauging the ‘wider impacts’ that the practice of rural development with special focus of “social mobilization” intends to determine how the living standard of the rural households is changing in communities that are participants of SRSO programmes for the last five years in the first five districts with ones that do not participate; in other words, what would happened if SRSO interventions were not made? Or more specifically: do the SRSO interventions have any impact on the lives of the poor? For this evaluation a quasi-experimental design (with baseline) of impact assessment with double-deference is applied. The baseline data was collected in 2004 and a followup survey in September 2008, includes 307 randomly selected households. The sample is stratified into control and treatment villages. The availability of baseline data for control and treatment groups would allow us using the “double difference” method for isolating the impact of SRSO on the participating households-that is a complex task, because there are many other factors that influence the outcomes and may not be the direct impact of SRSO. However the double differencing removes time invariant differences in factors influencing the outcome between the project and comparison groups. But, the validity of the double difference estimate still relies
upon the assumption that external determinants of the outcome were the same for treatment and comparison groups during the course of the intervention. (IEG World Bank). The findings of this study will be shared in the conference on 29 March – 2 April 2009 in Cairo.

| Themes 2: Impact evaluation: Approaches and methods | Session number: [35.3] |
| Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks | |
| Session stream: Rural development | Session name: Rural development 3 |

**Impact Evaluation of Poverty Alleviation in Nigeria: A Case Study of the National Poverty Eradication Programme (NAPAP)**

**Presenter: Laura Pan Luo**

Poverty has become a common feature of many developing countries, including Nigeria. The challenges that confront poor people can be summarized into lack of access to capital, information and stable markets. In Nigeria, the poor have limited access to capital and when they do, are faced with extremely high interest rates. Attempts were made by previous administrations to eradicate (or at least reduce) poverty through initiatives like Family Economic Advancement Programme (FEAP), Peoples Bank, Community Banks and Directorate for Food, Roads and Rural Infrastructure (DIFFRI). However, economic reforms can also initially exacerbate poor income distribution. To address this situation, the National Poverty Eradication Programme (NAPEP) was set up to enhance mass participation in economic development process, empowering the poor to have a voice, and therefore a way of expressing their ideas. Launched in January 2001, the program seeks to involve communities, cooperatives and individuals, by enhancing their capacity to become more productive. The NAPEP acts as the primary government agency for coordinating and monitoring all poverty eradication efforts at Federal, State and Local Government levels. It also assists the Federal Government in formulating poverty reduction policies nationwide, and intervenes in specific poverty reduction areas to provide social protection through economic empowerment as may be needed. It thus helps the government in creating mass participation in the development process in local communities. Because Nigeria is a federation, the structures for poverty alleviation are replicated in state and local government levels. Some of NAPEP's intervention schemes are the Youth Empowerment Scheme (YES), Rural Infrastructures Development Scheme (RIDS), Social Welfare Services Scheme (SOWESS), and Natural Resources Development and Conservation Scheme (NRDCS). These schemes have recorded among other things intervention by direct credit delivery, farmers empowerment, rural/urban transportation project, community skills development centres and collaboration activities with international development partners, notably the UNDP and World Bank. NAPEP has recorded some measure of success across the nation by 1) improving the socio-economic conditions of women, especially those in the rural areas through the provision of loan assistance, skills acquisition, reproductive health care services, adult literacy and girl child education; 2) building community capacities for wealth creation among enterprising poor people and promoting sustainable livelihood by strengthening rural responsive banking methodology; and 3) eradicating poverty through provision of microfinance and skill acquisition for income generation. However, there are still some obvious challenges to the poverty reduction scheme in Nigeria. These include poor coordination of activities, weak policy formulation, lack of sustainability of programmes and projects, absence of achievable target setting, poor monitoring, evaluation and impact assessment, duplication of functions, etc. There is need for proper monitoring of the various programmes of NAPEP with a view to addressing inherent weaknesses for a successful outcome.

Presenter: Victor Owusu

The emphasis over the years has been placed on agriculture through the development and diffusion of appropriate technologies. However, general increase in national food production does not necessarily ensure household and individual food security. In the economics literature, participation in non-farm employment has been found to improve household welfare in terms of the overall income and food security. Participation in non-agricultural activities allows farmers living near subsistence levels to acquire cash to meet their basic needs in addition to that supplied by own production. In most rural Ghana particularly in the Northern Regions, food insecurity and poverty are high. In other to survive the food deficits, most rural households in the region adopt various livelihood strategies. The paper examines the impact of non-farm employment activities on household food security. The rural farm household is food secured if it can satisfy its food needs without mortgaging standing field crops for current consumption. This paper employs a cross-sectional data collected from ten villages in Savelugu-Nanton District of Northern Region of Ghana. About 91 percent of the sampled households participated in non-farm employment activities which contributed on the average to about 60 percent of the total household income. The impact of non-farm employments on household food security was analyzed with the propensity score matching techniques where various matching algorithms were employed to estimate the average treatment effects. The data sets also allowed us to undertake separate estimations of the impact of husband’s non-farm employment activities and the impact of spouse’s non-farm employment activities on household food security. This approach is necessary since women in sub-Saharan Africa play a crucial role in ensuring that the household’s food needs are met. The results from the study indicate that nonfarm employment has a positive and robust effect on farm household income and a positive and significant effect on the likelihood of the household being food secured. It is therefore recommended that much effort be made towards the promotion of alternative income sources such as establishment of micro-enterprises in the rural economy as a means of supplementing the primary activity of the rural sector which is mainly agriculture (both crop and animal production).
Impact of a poor area development project on household welfare: Evidence from western China

**Presenter:** Ruchira Bhattamishra

**Co-author:** Luch Christiaensen

Although China was witnessed an impressive rate of poverty decline since economic reforms were implemented beginning in 1978, the declines have been uneven across provinces, resulting in geographic disparities. Given this, the Chinese government has implemented poverty-reduction programs specifically targeted at "poor areas". However, the existing evidence on the impact of these programs is mixed. In this study, we provide new evidence on the impact of poor area development programs in China. We examine the impact of the World Bank’s Western Poverty Reduction Project (WPRP) in Gansu and Inner Mongolia, two poor and lagging provinces in western China. The project was implemented between 1999 and 2004 and provided credit and technical assistance towards modern agricultural and animal husbandry techniques and also supported rural infrastructure.

Using a panel household survey collected in both project and non-project villages in the two provinces, and applying propensity score matching, we estimate the intent-to-treat effect of the WPRP on a number of indicators including per capita consumption, net income, business investment and gain output. We also provide preliminary evidence on the impact of the project on the potential pathways underlying poverty reduction by examining if it has differential impacts on intermediate outcomes (such as income from agricultural and non-agricultural sources). Clearly, matching provides valid program estimates only if project placement and participation are based on observable characteristics. However, we have access to a rich set of data on village and household characteristics at baseline that potentially explain program placement and participation, as well as outcomes of interest. In addition, the same survey instruments were implemented for households in both project villages and non-project villages, and both sets of households face the same economic and market conditions. As a result, we argue that matching is an appropriate method to generate a comparison group. In addition, since we have panel data before and after the project was implemented, difference-in-difference estimation permits us to eliminate bias due to unobservable, time-invariant differences between the treatment and comparison group not controlled by conditioning on observable pre-program characteristics.

We estimate project impacts for Gansu and Inner Mongolia separately since they had different project implementation agencies, received different project sub-components, and also had different pre-program characteristics (with Gansu being considerably poorer than IM). Our preliminary results indicate that the project had no impact on consumption expenditure, corroborating results from recent related studies. However, we find that the project has a positive impact on net income in Inner Mongolia which appears to be driven by an increase in agricultural income. This suggests that initial conditions may be critical in determining project impacts.
Evaluating Rural Electrification Programs: Reducing or Illuminating Poverty?

Presenter: Jorg Peters

This paper provides a practitioner-oriented guide to implementing state of the art evaluations in rural electrification programmes, following closely the work by Ravallion (2008a), Frondel and Schmidt (2005) or Angrist and Kruger (1999). It presents pragmatic possibilities to identify the counterfactual situation, taking into account specific challenges that researchers face when involved in such interventions. The focus is on projects that systematically provide electricity to a specific region – be it via grid extension, grid densification, or institutional support to public or private providers of decentralised electricity sources.

While evaluations ideally draw on panel data collected before, after, and maybe even during the intervention, this is the exception in most electrification projects so far. Researchers rather encounter either the situation after the intervention has ended with no appropriate baseline data, or before it is implemented. Furthermore, funds for evaluation research are very limited, making large or even country-wide surveys impossible in most cases. Therefore, in addition to the case of having panel data at hand, this paper offers options to derive robust insights on impacts of electrification using cross-sectional data of limited sample size collected either before or after the intervention (see also Ravallion and Wodon 1998, Kondo et al. 2008).

In particular, the paper argues in favour of examining impacts before the electrification intervention takes place. Results from such ex-ante impact assessments deliver insights for the project design, thereby reducing the gap between evaluation researchers and practitioners (Ravallion 2008b). In addition, the collected data can be used for robust ex-post evaluation if the opportunity to conduct additional surveying should arise. Experiences from combined baseline studies and ex-ante impact assessments in several African countries are outlined.

The pivotal point in conducting cross-sectional evaluation in electrification projects is to properly identify a control region. In the case of an ex-ante impact assessment, an electrified region that is comparable to the project’s target region has to be found. For this purpose, researchers familiar with the methodological requirements should visit at least some of the target villages and potential control regions and examine in cooperation with local professionals respective conditions according to a list of criteria.

Given that such comparable control regions can be identified, the paper demonstrates that robust evidence on poverty impacts can be derived even under tight budgetary restrictions and without before-after data. A crucial condition is to follow a mixed-methods approach that principally relies on quantitative data to credibly determine the counterfactual, but complemented by qualitative information from stakeholders in the project and key informants (White 2008). Not least, evaluation researchers should be in close contact with both project staff and the field research team. Having young researchers on the ground during the entire survey, turns out to be the most practicable set-up. This assures optimal compliance with methodological requirements and awareness about potential caveats that show up during implementation.

The holistic survey approach integrating quantitative data collection and indispensable qualitative information allows for both rigorous evaluation analysis and qualitative background information that can be fed back into intervention design.
Poverty Impacts of Rural Electrification Programmes – Evidence from Rwanda

Presenter: Jochen Kluve

Co-authors: Gunther Bensch, Jorg Peters

This paper presents practical experiences gained during the implementation of a baseline study conducted for an upcoming micro-hydro electrification project of German Technical Cooperation (GTZ) in rural Rwanda. Furthermore, findings from an impact analysis based on the collected data are presented, applying rigorous evaluation techniques.

The primary purpose of the baseline study is to document the socio-economic conditions in five villages in Rwanda prior to the implementation of the electrification project. In this project, GTZ assists five private operators both financially and technically in establishing and running mini-grids in rural areas. To allow for an ex-ante impact assessment, the survey design additionally includes a control group of households in twelve villages that had been electrified two to twenty years before the survey was implemented. We chose these control sites after examining potential villages throughout the country with regard to comparability with the project sites. Comparability indicators comprised the geographical region, market access, rural agricultural conditions, and access to infrastructure. This control group allows us to assess how households and microenterprises in the project region will perform in the wake of electrification.

In total, using structured questionnaires 735 households were interviewed, 400 in the target and 335 in the control region. In addition, a smaller, more qualitative survey was conducted among some 200 micro-enterprises and social institutions. To complement these quantitative and semi-quantitative data, key informants like local chiefs, NGOs, or project staff were also interviewed. Since one major aim of the study was to gain insights for the electrification project, the researchers were in close contact with the project staff and results were fed back into the intervention design.

For the purpose of the impact assessment, we first investigate the data descriptively with regards to intermediate outputs and central development outcomes. The major outcome variables in our analysis are income, educational and health indicators. On the level of intermediate outputs behavioural changes can be detected reflecting a substantial usage of the provided service. E.g., electrified households consume 350 times more light in terms of Lumen hours than households relying solely on traditional lighting sources. This, however, does not translate into measurable impacts on the level of major development outcomes such as income.

In addition, statistical matching techniques are applied to further increase comparability of electrified and non-electrified households. For this purpose, we apply propensity score matching using household variables that are considered to be less sensitive to the electrification intervention (e.g. educational attainment of the household head, construction material of the dwelling, and food expenditures).

In fact, the findings from this matching approach reveal several statistically significant outcomes and are partly supportive of existing anecdotal evidence on electrification impacts. One notable result is that school children exhibit considerably higher home study hours in electrified households than in comparable non-electrified ones. Results concerning income suggest a positive or ambiguous impact depending on the examined indicator. Hence, addressing the selection into treatment that is usually strong in electrification projects via matching techniques seems to improve the quality of the impact assessment.
The Welfare Impact of Rural Electrification

**Presenter: Howard White**

Many claims are made regarding the welfare impacts of rural electrification, ranging from improved education and health impacts to positive impacts on income through enterprise development. New analysis by IEG finds empirical evidence for many of these links. Bank rural electrification projects have become more explicitly focused on poverty reduction. However, beyond increase in electricity supply, complementary measures to ensure the highest poverty reduction impact – like educational campaigns, promotion of productive use and smart subsidies – have been generally lacking. While Bank projects have generally succeeded in establishing infrastructure for electricity transmission and distribution in rural areas, the finances of many providers remain weak as the cost of supply is not met by revenue.

However, this need not be as consumer willingness to pay is almost always at or higher than cost recovery rates for suppliers- though advances need to be made in reducing system losses and overcoming the high connection charges that are a barrier to potential consumers.

An ex-ante evaluation of a conditional cash transfer for children in Uruguay

**Presenter: Veronica Amarante**

**Co-authors: Rodrigo Arim, Gioia de Melo, Andrea Vigorito**

During 2007, an old child allowance program of Uruguay, named Asignaciones Familiares, was redesigned. The new design of the program deepens the non contributory nature of the regime, extending significantly the coverage of the social protection network among children from the first income quintile. Its main aims are increasing household income and fostering attendance to the educational system at secondary school.

This research aims at shedding light on the effects of this renewed conditional cash transfer program on teenage school attendance and work, adult labour participation, poverty and income inequality, by carrying out an ex-ante evaluation. This exercise can provide inputs for policy implementation and further improvement of the intervention. Our work is based on the methodology developed by Bourguignon et al (2002) for Bolsa Escola.

We focus on the impact of the policy change on teenage child schooling, poverty and inequality, considering that the program only affects household well-being through child behaviour. We develop microsimulations on teenage schooling behaviour, considering five scenarios. The first three options correspond to the results of the microsimulation exercise based on the characteristics of the new regime and two additional schemes with the same total cost. To analyse the impacts of the new transfer on poverty and inequality, we consider two additional
extreme scenarios. The fourth one assumes that child schooling remains the same that at present, and hence, the policy exerts no effect on attendance. The fifth scenario assumes that, as a consequence of the transfer, all children attend school. This can be interpreted either as the removal of the conditionality of the transfer or as a full take-up scenario.

We also address the effect of the targeting instrument on the results obtained. In the previous literature on ex-ante evaluations, the interferences generated by the use of proxy-means tests to target beneficiaries was not addressed, as most of them are based on the target population defined on the basis of observed income in household surveys. In this research we compare the effects of the program on the ideal target population with the ones obtained when the proxy means score is used to select beneficiaries. By computing the proxy means score, we are able to analyze to which extent the targeting procedure entails losses in the ability of the program to reduce poverty and inequality.

Finally, we explore the effects of the new program on adult labour supply. In order to microsimulate behavioral responses of adult labor supply we estimate a labor supply model for adults, and we use the corresponding parameters to simulate the effect of changes in cash transfers under the three scenarios considered. The exercise provides useful information about what to expect in terms of adult labor market behavioral responses to cash transfers and sets questions for further research.

### Themes 2: Impact evaluation: Approaches and methods

### Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks

<table>
<thead>
<tr>
<th>Session stream: Transfer Payments</th>
<th>Session name: Transfer Payments 1</th>
</tr>
</thead>
</table>

**Beyond Cash: Assessing Externality and Behavioural Change Effects of a non-randomized CCT programme**

**Presenter:** Rafael Perez Ribas

**Co-authors:** Fábio Veras Soares, Guilherme Issamu Hirata

Impact evaluations of Conditional Cash Transfer (CCT) programmes usually assess both core objectives and possible unintended effects. Several studies have shown that these programmes tend to affect decisions on time and budget allocations, mainly in favour of children. Nonetheless, it is unclear whether these changes stem from the income increases or from other components of CCT programmes that aim to change the behaviour of the families. If these impacts were mainly explained by the relaxation of the budget constraint, then the other components of these programmes might represent an unnecessary cost. But if the monetary transfer were not enough to induce desired changes, other components would be relevant.

Another important issue is the impact of externality effects on potential outcomes of beneficiaries. Due to social interaction, households can be affected by the mere existence of beneficiaries in their community, whether or not they themselves are participating into the programme. This process affects the outcomes of both beneficiaries and non-beneficiaries and can either increase or attenuate the potential impact of the programme. Failure to take externality into consideration can lead to biased estimates of the average effect of the programme on the treated.

To disentangle the contribution of each component of a CCT programme, one should ideally carry out an experiment that allows for multiple treatment design. This could be obtained.
randomizing the distribution of each component. Similarly, to estimate externality effects one should apply a random, two-level strategy of assignment. The first level consists of an experiment comparing communities. The second level consists of an experiment comparing households within communities. However, both kinds of experiment may be difficult to implement, even in pilot phase.

This paper presents a quasi-experimental method applied to the consumption data from Tekoporã (Paraguayan CCT) impact evaluation. The presence of two potential comparison groups, within-and between-communities, allow us to disentangle the programme’s average effect on the treated into participation effect (being a beneficiary) and externality effect (being in a community where there are beneficiaries). In addition, the method allows for these effects to be further decomposed into “income effect” and “behavioural change” (other components) effect.

With regard to the results, the total impact on per capita consumption has been negative, despite the positive effect of both participation components—income and behaviour change effects. The negative result is completely due to the externality effect, possibly derived from social interactions among households. By the same token, most of the total positive effect on household saving is explained by externality.

Tekoporã has also had a negative impact on food share, mainly due to participation effect caused by behaviour change. On child-clothing share, there have been neither income nor externality effects; the positive impact stems entirely from behaviour change caused by the programme’s other components.

Distinct outcomes are thus affected through different mechanisms. Therefore, managers should know which of them are more effective and efficient for the purposes of meeting programme goals, and through with channels they work, opening up the black box of the impact evaluations.

**Themes 2: Impact evaluation: Approaches and methods**

**Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks**

**Session stream: Transfer Payments**

**Session number: [37.1]**

**Session name: Transfer Payments 1**

**Investing Cash Transfers to Raise Long Term Living Standards**

**Presenter: Sebastian Martinez**

**Co-author: Gertler and Rubio**

Using data from a randomized experiment, we find that poor rural Mexican households invested part of their cash transfers from the OPORTUNIDADES program in income generating micro-enterprise and agricultural activities. For each peso transferred, beneficiary households used 88 cents to purchase consumption goods and services; and invested the rest. The investment yields a 1.8-cent increase in long-term consumption for every peso transferred. This translates into a 7 percent increase in consumption after a year on the program, a 32 percent increase after 5.5 years (the period observed in the sample) and 47 percent after nine years -- the minimum period of benefits. These results suggest that cash transfers to the poor can raise long term living standards through investment, which if maintained after program benefits end can contribute to long term poverty reduction.
Labor supply responses to cash transfer programs: Experimental and non-experimental evidence from Latin America

Presenter: Guillermo Cruces

Co-authors: María Laura Alzúa, Laura Ripani

This document analyzes the impact of conditional cash transfers programs on the decision to work in a series of Latin American countries. The rapid expansion of this type of programs throughout the region has spurred a wealth of empirical studies, but the labor supply of adults has been seldom covered in this literature. This document aims to fill this gap. It presents results from the experimental evaluation of programs in Mexico, Honduras and Nicaragua, and results from non-experimental evaluations from Mexico, Brazil and Colombia. The paper also discusses the implication of different evaluation methods, ranging from experimental approaches (conditional differences in differences) and non experimental methods (propensity score matching).

Conditional cash transfer programs (CCTs) have been hailed as a new paradigm in social policy in Latin America, and their influence has spread beyond the region. The popularity of this type of initiatives can be partially traced to the implementation of an original and well-executed experimental impact evaluation methodology for the PROGRESA program (now Oportunidades) in rural Mexico in the mid 1990s (see Parker and Skoufias, 2000). The positive impacts of this program on the targeted outcomes, obtained through a credible methodology, helped the dissemination of CCTs and of experimental evaluation methods.

The rapid expansion of programs of this type throughout the region has spurred a wealth of empirical studies. There is a vast literature on the impact of CCTs on the programs’ main target outcomes: educational and health results, nutrition, household income, and child labor (see for instance the review of Rawlings and Rubio, 2003). However, the labor supply of adults has received little attention in this literature. The first contribution of this document is to expand this line of research, complementing this previous work for PROGRESA and providing a series of comparable empirical results from experimental evaluations of other CCTs in Latin America.

While human capital accumulation is the main long term goal of CCTs, from a policy perspective it is still important to understand any impacts of the cash transfer component on the decision to work. The existence of any impact, and its direction, is ultimately an empirical question, since standard labor economic theory does not provide an unambiguous prediction for the labor supply response to an increase in unearned income. The empirical work presented also allows the comparison of results from experimental and non-experimental evaluation methods.

The document discusses the theoretical underpinnings of the impact of cash transfers on labor supply, and presents a review of the empirical literature for countries in Latin America. It also discusses the evidence from three programs that implemented an experimental evaluation strategy: PROGRESA from Mexico, Red de Protección Social from Nicaragua and PRAF from Honduras. These experimental evaluation strategies, however, have been the exception in the region. The document also presents impact-evaluation studies based on propensity score matching techniques for Brazil, Colombia and Mexico. The document concludes by discussing the implications of experimental and non-experimental evaluation techniques.
Designing quasi-experimental impact evaluations of Cash Transfer programmes: IPC experience with Mozambique and Dominican Republic cash transfer programmes.

**Presenter: Fabio Veras Soares**

**Co-authors: Rafael Perez Ribas, Guilherme Hirata**

This presentation summarizes the main challenges around the design of quasi-experimental evaluations of cash transfers in developing countries.

A common trait of these evaluations is that the programmes did not foresee and/or plan an impact evaluation from the outset. In most cases, the demand for an impact evaluation came when the programme was already being implemented. Despite the well known case of the experimental evaluation of Progresa in Mexico, a quasi-experimental setting is actually the rule and not the exception for this type of programme in developing countries. For this reason, the dissemination of the experiences on how to deal with the usual (and sometimes not so usual) challenges posed by this type of design is extremely important for the consolidation of best-practices in this area. Such knowledge would allow both programme managers and researchers to learn from the mistakes and achievements from other experiences and improve the design of future evaluations.

The International Poverty Centre – IPC – has recently being involved in the design of the impact evaluation of two cash transfer programmes that were already working for some time: Solidaridad (Solidarity) in Dominican Republic and Programa de Subsidios de Alimentos - Food Subsidy Programme in Mozambique. Different strategies and sampling options were tailor-made for each case. These choices were based on the design of the programme and on the availability of administrative information to inform the sampling. In this presentation, we discuss for each of these programmes the following issues:

a) the use of the expansion of the programme into new areas as means to have a proper baseline and to try to replicate an experimental setting.

b) the challenge of synchronizing baseline survey and incorporation of new beneficiaries in the programme;

c) how the comparison group was defined;

d) the importance of administrative records to inform sampling;

e) the problem of not knowing the population from which one has to draw the sample – how to weight observations;

f) how to choose the appropriate quasi-experimental techniques: selection on observables – propensity score matching and regression discontinuity design (oversampling around the cut-off point) – and selection on unobservables (the hard task to look for instrumental variables) and differences-in-differences (when a baseline is available); and

g) how to prepare the questionnaire: outcome and controlling variables.

Design options that worked and those that did not work will be discussed to highlight what one can do to preserve the evaluation when things go differently from what was originally planned.
Achievements and Shortfalls of Conditional Cash Transfers: Impact Evaluation of Paraguay’s Tekoporã Programme

**Presenter:** Fabio Veras Soares

**Co-authors:** Rafael Perez Ribas, Guilherme Hirata

Tekoporã is a Conditional Cash Transfer (CCT) programme in Paraguay. The programme seeks to reduce extreme poverty by using direct cash transfers to poor households with children and diminish the potential for future poverty by encouraging investment in human and social capital. The programme intends to break the intergenerational cycle of poverty through investments in the health and education of children.

Households are eligible for the programme if they fulfill all the following conditions: 1) have children under 15 years of age or pregnant women; 2) live in the priority areas of the programme, namely, the poorest districts in the country; and 3) have a low Index of Quality of Life (ICV). ICV is a non-monetary measure varying from 0 to 100 that synthesises several quality of life dimensions, such as access to public services, health and education outcomes, occupation of the household head, housing conditions and household assets. The programme provides the transfer to those households that are classified as extremely poor (having an ICV below 25) or moderately poor (having an ICV between 25 and 40). The transfers are conditional on school attendance, regular visits to health centres and updating of immunizations. Tekoporã also includes a family support initiative that, among having other effects, should increase the productive potential of the household and its social participation.

The evaluation of a pilot project can offer important inputs into the decision-making process on the feasibility of the scaling-up of the programme, the effectiveness of its design and the assessment of problems that limit its potential. This paper presents a summary of the impact evaluation of the pilot of Tekoporã. We start presenting the database available for the evaluation. Then, there is a discussion on the rationale for building appropriate comparison groups based on untreated observations, as well as the evaluation strategy, i.e., how we tried to circumvent the so-called ‘fundamental evaluation problem’. Special emphasis is given to the assumptions used in the definition of each estimator implemented in this evaluation, namely, the Difference-in-Differences and the cross-sectional average difference, and the Propensity Score techniques that were adopted to ensure that both treated and untreated households were similar.

The evaluation shows positive impacts on per capita income and consumption, poverty reduction, school attendance, investment in agricultural production, access to credit, savings and social participation. Indeed, these results easily justify its scaling-up. However, the pilot has not been successful in reducing child labour or increasing child immunizations. Thus, addressing these aspects needs to be a key part of any redesign of the programme when it is scaled up.
Heterogenous impact of the social program Oportunidades on contraceptive methods use in young adult women living in rural areas: limitations of the regression discontinuity approach

**Presenter: Hector Lamadrid Figueroa**

Oportunidades is a social program run by the Mexican government seeking to improve education, health, nutrition, and living conditions of those in extreme poverty. People supported by the program attend monthly health talks, which include information on contraceptive methods. Reduction in fertility, especially among youths, is deemed crucial to accomplish the program’s goals. The government defined as eligible to become enrolled those below a certain threshold of a poverty index or score, calculated at the beginning of the program, the index had a mean and standard deviation of 743 and 144 points respectively, and the threshold for eligibility was 752 points. WE analyzed information on contraceptive methods use in a sample of 2230 young women from the Oportunidades evaluation surveys conducted in the years 1997 to 2000.

Among different approaches, the program’s impact was evaluated by Regression Discontinuity Analysis comparing the prevalence of contraceptive methods use among those who were just above or just below the poverty score cutoff point for eligibility. We compared treatment area subjects who were just above vs. those who were just below the cutoff point at windows of 50, 100, and 150 points. This strategy allowed us to compare eligible vs. non-eligible subjects who have nearly the same poverty, so that being eligible is their only difference. Models were fitted by means of both Ordinary Least Squares (OLS) and Two-Stage Least Squares (2SLS) adjusting for poverty, age and state, with robust standard errors and clustering at the locality level. In contrast to what we found in Intention to Treat analysis (5-10 percent points increase), results showed large, negative, and statistically significant program effect estimates.

One possible explanation for these results is that the program does have an effect but it is not constant relative to poverty, being greater on those who are very poor and being deleterious at poverty levels near the eligibility threshold. To test this hypothesis, we estimate, by means of locally weighted scatterplot smoothing (LOWESS), the relationship between the poverty index and the proportion of contraceptive methods by treatment area, above and below the cutoff point for eligibility. Using this approach it became clear that the strongest effect of the program is on the poorest of the poor, being close to 30 per cent points higher in those in Oportunidades areas below a poverty score of 600, as opposed to being nearly zero in the vicinity of the cutoff point for eligibility. The relationship between poverty and contraceptive methods use for those not eligible for the program in similar for treatment and control areas, as would be expected. The unexplained drop in the prevalence of contraceptive methods use just below the cutoff point is responsible for the apparently negative effect of the program, and gives the impression of a strong discontinuity in an sense opposite to that expected.

In consideration, this is an example were Regression Discontinuity can fail to find an impact of the program if the main impact occurs far from the eligibility threshold, even if the impact is overall large.
Pensions, Poverty and Household Investments in Bolivia

**Presenter: Sebastian Martinez**

This study estimates the effect of the BONOSOL pension to elderly Bolivians on household consumption and investments. Pension eligible households in rural areas increase food consumption by two times the value of the transfer. Even if the marginal propensity to consume food was equal to one, this result indicates that households multiply the transfer to obtain a greater aggregate benefit than if the money had been spent directly on final goods. Rural beneficiary households in Bolivia, in possession of land assets but with limited access to monetary capital, obtain multipliers on the transfer through investments in agricultural production. In support of this argument, approximately half of the increase in food consumption is derived from home production of meats, vegetables, and other agricultural products. Additional evidence on farm investments, land use and animal ownership supports the hypothesis that beneficiary households use the pension to boost consumption through more home production. When considering the use of cash transfer programs for poverty reduction, governments and multilateral institutions should consider the additional benefits obtained by beneficiaries who put their transfer money to work.

The Heterogeneous Impact of CCT Programmes on Child Labour – The case of Tekoporã in Paraguay

**Presenter: Guilherme Issamu Hirata**

**Co-authors: Rafael Perez Ribas, Fábio Veras Soares**

One of the major challenges faced by developing countries in their fight against poverty is to break the intergenerational cycle of poverty. In order to guarantee a minimal income level, poor families often encourage their children to work. As a consequence poor children have less time to spend studying than non-poor children. These children end up with a lower educational achievement which reduces their capacity to generate income in the future.

Condition Cash Transfer (CCT) programmes are policy interventions whose main goal is to break that cycle through incentives given to the families to improve their children’s health and education conditions. In this context, the conditionalities attached to the cash benefits have the purpose to increase the opportunity cost of sending children to work.

However, child labour is not usually listed among the programme’s objectives and specific components of the programme design are not implemented to address this issue. In spite of
This, it is always expected that CCT programmes have effects on child labour due to their direct impact on child time allocation, which also results in an indirect impact on school attainment. In this regard, CCT programme evaluation constitutes a great opportunity to improve our understanding on the determinants of child labour, as long as it generates an exogenous shock to the living conditions of poor families, which may change their attitude towards child labour.

This paper aims to contribute to the literature of the determinants of child labour using a CCT programme as an object of study. The idea is to exploit the evaluation survey of the Paraguay’s CCT Pilot Programme, called Tekoporá, to investigate both the direct impact on child labour and also some of the mechanisms that led to that result. In this sense, besides estimating the Average Treatment Effect on the Treated (ATT), we will also analyze the heterogeneity of that impact to shed some light on the household decision making process regarding labour supply choices.

First, we investigate whether the cash transfer, jointly with its conditionalities, has a direct effect on child labor decision and whether this effect is heterogeneous according to children's birth order. The hypothesis of heterogeneity is based on the Rotten Kid Theorem (Becker, 1974). Second, we analyze whether child labor supply increases along with parents’ labor supply. The results show that birth order is important when the child decision of working is not jointly taken with her parents. When we consider the joint decision, this effect disappears. We also found that the child labor is pro-cyclical to the mother’s supply, regardless the father’s decision. However, it seems that the child labor is negatively related to the father’s supply, but only indirectly, through the negative impact that the father’s labor supply choice has on the mother’s labor supply.

<table>
<thead>
<tr>
<th>Themes 2: Impact evaluation: Approaches and methods</th>
<th>Session number: [37.4 ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks</td>
<td>Session stream: Transfer Payments</td>
</tr>
</tbody>
</table>


Presenter: Pedro Olinto

Conditional Cash Transfer (CCT) programs aim to reduce poverty by making cash payments to poor households conditional on investments in the human capital of children. Conditions usually include enrolling children into public schools, getting regular check-ups at the doctor's office, receiving vaccinations, or the like. Countries have been adopting or considering adopting conditional cash transfer (CCTs) programs at a prodigious rate. In some countries - including Mexico, Brazil, and Ecuador - CCTs have become the largest social assistance program, covering millions of households. They have been hailed as a way of reducing inequality, especially in the very unequal countries in Latin America; helping households break out of a vicious cycle whereby poverty is transmitted from one generation to another; promoting child health, nutrition, and schooling, and helping countries meet the Millennium Development Goals.

Do these and other claims make sense? Are they supported by the available empirical evidence? What about countries that do not have CCTs but are considering implementing them, often in circumstances very different from those where they were first introduced?

A new world Bank report, Conditional Cash Transfers: Reducing Present and Future Poverty, Policy Research Report, seeks to answer these and other related questions. The report lays out a conceptual framework for thinking about the economic rationale for CCTs;
reviews the very rich evidence that has accumulated on CCTs; discusses how the conceptual framework and the evidence on impacts should inform the design of CCT programs in practice; and discusses how CCTs fit in the context of broader social policies.

The authors show that there is considerable evidence that CCTs have improved the lives of poor people and argue that conditional cash transfers have been an effective way of reducing poverty in the short run and contributing to its reduction in the long term through investments in human capital of poor children. They also recognize that even the best-designed and managed CCT cannot fulfill all of the needs of a comprehensive social policy. They therefore need to be complemented by other interventions, including those oriented to improving the supply of health and education services.

### Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks

<table>
<thead>
<tr>
<th>Session number: [38]</th>
</tr>
</thead>
</table>

**Session stream: Vocational training and business support services**

**Session name: Vocational training and business support services**

---

**Evaluating the Impact of Job Training Programs in Latin America: Evidence from IDB funded operations**

*Presenter: Pablo Ibarraran*

*Co-author: David Rosas Shady*

Among active labour market programs, job training is popular in Latin America as an attempt to help the labour market insertion of disadvantaged youth, and also as a way of providing skills to low-income groups to enable them to deal with the challenges of globalization. This paper summarizes the findings from the first rigorous set of evaluations to job training programs in Latin America that were made in the context of a project undertaken by the Office of Evaluation and Oversight at the Inter-American Development Bank. This research was complemented by two independent impact evaluations of similar training programs in Chile and Colombia. We report the results of two evaluations with an experimental design (the Dominican Republic and Colombia), one with a natural experiment (Panama) and four non-experimental evaluations (Argentina, Chile, Peru and Mexico). Overall, the results suggest that employment effects range from modest to meaningful—increasing the employment rate by about 0 to 5 percentage points—although higher and significant for some groups such, as women in Colombia and Panama—with impact of 6 to 12 percentage points in the employment rate. In most cases there is a larger and significant impact on job quality, measured by getting a formal job, having a contract and/or receiving health insurance as a benefit. Finally, we present an operational definition of the impact of training on “employability” in the context of a dynamic model with state dependence and unobserved heterogeneity, which we were able to apply in the evaluations of the Dominican Republic, Panama and Argentina.
Impact assessment of Learnerships and Apprenticeships for the Manufacturing, Engineering and Related Services Seta (MERSETA), South Africa

Presenter: Jeffy Mukora

Co-author: Mariette Visser

This paper will present the methodology used and findings of a study that was commissioned by the MERSETA with the aim of ascertaining the efficiency and effectiveness of the learnership and apprenticeship systems and to assess their impact on the demand for and supply of skills for the industry. The aims of this research were influenced by the fact that an understanding of the potential impact of learnerships and apprenticeships on the labour market outcomes of beneficiaries within the sector was limited. There was a critical lack of data on the scale, number and career progression of qualified apprentices and learners as well as the employability of newly qualified learners exiting at different levels of the National Qualifications Framework (NQF). The status of many participants, or their motives for studying and moving within the system, or the different possible pathways open to them or that they traverse was not known. Such information which must be as reliable and accurate as possible was needed in order to enhance MERSETA’s ability to strategically intervene in training initiatives geared towards addressing the supply of and demand for skilled labour within the sector.

The research deployed both quantitative and qualitative methods. The quantitative analysis was based on datasets provided by the MERSETA. The databases consisted of the total population of participants in the learnership and apprenticeship systems within the sector.

The analysis of the datasets reflected the shape of the learnership and apprenticeship systems in terms of NQF levels, trades and qualifications as well as the demographic profile of the total learnership and apprenticeship population participants (gender, race, age, disability, geographical distribution). The datasets also provided telephonic contact details for learnership and apprenticeship participants and formed the population from which random samples for the survey was drawn.

Qualitative data were produced through interviews with representatives from the majority of the stakeholder groups involved in the learnerships and apprenticeships for the MERSETA at the levels of provision and policy. In addition five provinces were visited and a sample of training providers and employers were interviewed. In-depth interviews with learners traversing specific pathways related to Learnership and Apprenticeship programmes in the low, intermediate and high skills bands were undertaken. An extensive review of existing data, policy documents and relevant legislation was also undertaken as part of the methodology.
Introduction. Small medium and micro enterprises (SMMEs) have potential to stimulate local economic development, through employment creation and the promotion of sustainable livelihoods. Over the years stakeholders have realised the need to capitalise on this potential through market development approaches. Umsobomvu Youth Fund Voucher programme provides technical and managerial support to SMMEs to increase entrepreneurship amongst young people in South Africa. An evaluation of the Business Consultancy Services (BCS) voucher was conducted to measure the impact of the voucher programme.

Methodology. The evaluation methodology followed a literature and document review to get an assessment of the South African BDS context, followed by fieldwork assessment of the UYF BCS Voucher programme implementation, and an assessment of the programme impact evaluation. In addition to this, the evaluation covered a review of the InTouch database and conducted interviews with all the stakeholders in the programme, these include: Allocating agents; Service providers; beneficiaries of the BCS Voucher services; and UYF staff working on the programme.

Findings. The programme is currently operational in 9 provinces working with 10 accredited third party organisations, 267 Service Providers and 22,616 Young Entrepreneurs. 42 services are accessible in 14 broad categories of tendering, web-based marketing, accounting, bookkeeping, business feasibility, branding, memorandum of understanding, legal services, business and financial administration, business registration and marketing plan.

Vouchers with the highest frequency are for business plans (58.1%), branding (9.1%), business registration (7.4%) and web-based marketing (4.2%). This pattern is consistent with the maturity of the programme and the target market which is dominated by new entrants in the business environment. The programme has both rural and urban outreach with urban areas receiving the highest number of vouchers. Provincial level analysis shows Gauteng being the largest recipient with 26.1% of the issues vouchers while the lowest is in Mpumalanga at 7%. There is a satisfaction with the process and results show efficient processing time of 1 week in the majority of cases. Approval of application is followed by payment of part fees by the Young Entrepreneurs which allows them to select and use any of the approved Service Providers in their area. Service Providers are selected based on recommendations from Third party organisations, consulting the service directory, recommendations from other clients, service offering, reputation and convenience in location. The evaluation found that there was a general satisfaction with the Service Providers considering value for money (85%), venue (93%), duration of service (74%) and the content delivered (86%).

The programme has also contributed to the establishment of enterprises with the majority, 70%, of the firms having been formed during the period when the programme was introduced. A relatively balanced gender outreach has been attained with 54.2% men and 45.5% women.

Recommendations include (i) improve the Information Flow between all stakeholders; (ii) Improve marketing of the programme using innovative methods; (iii) Quality control of Service providers to be improved at all levels; and (iv) Enhance performance management indicators for measuring impact.
### Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks

<table>
<thead>
<tr>
<th>Session number: [38]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Session stream:</strong> Vocational training and business support services</td>
</tr>
<tr>
<td><strong>Session name:</strong> Vocational training and business support services</td>
</tr>
</tbody>
</table>

**Laying Foundation for Impact Assessment: Introduction of Record Keeping, Costing and other data utilization culture to the rural and urban low income entrepreneurs**

**Presenter:** Samuel Migui Wachira

The objective of the study is to enable thirty self-help, community based organization (about 900 to 1,200 people) among the rural small scale farmers and urban slum dwellers efficiently and effectively keep records be able to use data in decision making and cost their transactions at the organizational and individual levels within a period of one year.

**Purpose of the study.** To enable measurement of economic impact of various economic development initiatives that are undertaken at this socioeconomic stratum of society.

**Hypotheses.** We start here with two hypotheses:

- Economic impact of self- help community based organizations now operating with renewed dynamism in Kenya will in the near future emerge as a major avenue of democratization of development and wealth creation. Hence calling for increased attention.

- One of the best ways to address this paradigm is through Development Evaluation approach where an evaluator involves self in operation of such groups at all stages of their operations, starting from planning through execution to final evaluation, would be of major impact particularly where the evaluator works with the respective group to address felt needs particularly in the area of generation and use of relevant knowledge.

- The knowledge of Record keeping and costing of enterprises of which this socioeconomic category is denied access would be a major development catalyst if available training models, comprehensible to this social category were democratized.

### Themes 2: Impact evaluation: Approaches and methods

<table>
<thead>
<tr>
<th>Session number: [39]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Session stream:</strong> Water for agricultural use</td>
</tr>
<tr>
<td><strong>Session name:</strong> Water for agricultural use</td>
</tr>
</tbody>
</table>

**Using social network analysis to evaluate the effect of Egypt’s water policy network characteristics on the vulnerability of its water system**

**Presenter:** Mohamed A. Hamouda

This study employs social network analysis (SNA) in investigating the links between water policy network characteristics and vulnerabilities of water systems. The investigation included outlining the processes of water policy formulation and implementation in Egypt, as well as, characterizing the interactions between the actors in the policy network. The relationship analysis among the actors in the Egyptian water policy making process focused on characteristics that define its effectiveness and inclusiveness. The study was designed to develop indicators of vulnerability in the water policy context using both quantitative and qualitative information on the arena of water policy making. The resulting developed indicators concentrate on evaluating certain relationships that are thought to weaken (or strengthen) the resulting policies. The developed indicators directly capture institutional factors that affect the resilience of Egypt’s water system.
Methodology. A list of actors in the policy making process and their respective roles was prepared based on experts’ opinions. A questionnaire investigating the relationships among these actors was designed and a survey was carried out to gather data on relationships among actors within the network of water policy making. The quantitative analysis was done using UCINET, a software for Social Network Analysis. The results were integrated with the qualitative analysis of the network to give an idea on the vulnerability of the water policy network in Egypt.

Results. The case evaluation for Egypt indicated that the conducted Social Network Analysis met the necessary standards for analytical soundness and that the output contained several insights that were considered to be credible and relevant by the involved water experts. Water policy indicators were evaluated and underlying characteristics of water policy vulnerability/resilience were identified, these were primarily related to the weak role of non-governmental organizations, independent research institutions, and water users associations; as well as, the generally low level of effective relationships within the network. On the other hand several recommendations for alleviating the detected deficiencies in the water policy network were outlined. The developed indicators can be a valuable addition to the wealth of indicators evaluating the vulnerability of water systems provided by previous studies.

Themes 2: Impact evaluation: Approaches and methods

<table>
<thead>
<tr>
<th>Session stream: Water for agricultural use</th>
<th>Session number: [39]</th>
</tr>
</thead>
</table>

Participatory Poverty and Vulnerability Assessment for Appropriate Targeting and Impact Evaluation: The Case of the Lower Usuthu Smallholder Irrigation Project in Swaziland.

Presenter: Lynn Kota

The Lower Usuthu Smallholder Irrigation Project is an 8 year, US$190 million project co-financed by the International Fund for Agricultural Development (IFAD), European Commission (EC), European Investment Bank (EIB), Arab Bank for Economic Development of Southern Africa (DBSA), International Development and Cooperation Fund (ICDF) and the Government of Swaziland (GoS).

The goal of the project is the reduction of poverty and sustained improvement in the standard of living of the population in the Project Area through the commercialization and intensification of agriculture. The two key stated purposes of the project are:

a) The integration of smallholder farmers into the commercial economy through the provision of irrigation infrastructure, development of the policy and legal framework for smallholder irrigation as well as the establishment of farmer-managed irrigation institutions.

b) Sustainable improvement in the environmental health in the Project Area to ensure that the population derives the full benefit of agricultural commercialization.
Integration of the Most Significant Change Approach in measuring impact for market development projects

Presenter: Nyamwaya Munthali

Impact evaluation as the component of a project cycle that seeks to validate to what extent changes occurring to beneficiaries can be attributed to project interventions, is a topical issue in the development sector. Whether from a private or public organisation perspective impact evaluation is necessary in project planning and improving implementation strategies. As monitoring and evaluation practitioners in the developing world, where the race to meet the millennium development goals is on, the challenge is always how to innovatively and cost effectively use and adapt the basic approaches and methods of impact evaluation (quantitative and qualitative) for effective monitoring and project evaluation.

As a qualitative-participatory approach the (Most significant change) MSC model was introduced by the Mennonite Economic Development Association (MEDA) to all its projects to build on the core evaluation technique that falls on the conventional control and treatment group approach (baseline/follow up). MSC is an approach than borrows largely from the profession of journalism dwelling on immediate feedback, direct answers from the horse's mouth and use of stakeholder panels to dissect interventions that they appreciate or do not.

In Zambia, the MSC methodology is being applied on one of MEDA’s market development project that seeks to accelerate the supply chains of appropriate water technologies for smallholder farmers using a smart subsidy. The project has market facilitation activities on both supply and demand sides. The project also works with partners which serve as the avenue to access smallholder farmers. At project inception, all partners failed to avail baseline data for the targeted smallholder farmers. The MSC approach was introduced to supplement other research such as detailed individual impact case studies that served as a small “baseline survey” and a client satisfaction survey to measure the program impact on the targeted farmers as well as market development indicators. Among the key lessons on effectiveness of the MSC, particularly in this integrated monitoring and evaluation framework of the Zambia program and on other MEDA projects, are the following:

- It is holistic and covers a broad range of factors, discovers unanticipated positive and negative effects
- Gives a richer description of the context in which a positive or negative change is occurring.
- There is closer contact with stakeholders unlike in surveys
- Involves stakeholders leaving them to interpret the most significant change themselves, this adding an aspect of quality to the data analysis and determines effectiveness as affected parties bring out the relevance of intervention to them
- Provides instant feedback to validate quantitative analysis, MSC can be conducted at any time during the life of a project
- There is room to reflect more directly on perception gaps
- Achieves data analysis and interpretation at one go
- It is more flexible as it is not based on predefined indicators
This paper discusses the technical aspects of integrating the MSC in a monitoring and evaluation system of a market development program. The paper will argue that in the absence of baseline data, use of the MSC, together with other quantitative and qualitative tools serve as an innovative way of measuring impact. The paper will also demonstrate how these research tools can be integrated. Finally the paper will also present some of the initial results pointing to the benefits of this integrated approach.

<table>
<thead>
<tr>
<th>Themes 2: Impact evaluation: Approaches and methods</th>
<th>Session number: [39]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session stream: Water for agricultural use</td>
<td>Session name: Water for agricultural use</td>
</tr>
</tbody>
</table>

**Working Together to Enable Communities in the Aral Sea Basin to Combat Land and Water Resource Degradation**

**Presenter: Ikbal Yusupova**

IWMI has lead an ADB funded project entitled “Enabling farming communities in the Aral Sea basin to reverse water and land degradation by the creation of 'Bright Spots’” to address the overexploitation of water and soil resources in the region which has resulted in serious economic, social and environmental consequences.

The creation of Learning Alliances(LAs) was an important approach that has been adopted, aimed at facilitating the vital learning and sharing of knowledge necessary to developing and outscaling ‘bright spots’ practices. The Galaba Farmers LA, composed of farmers, specialists, local decision makers, and other representatives, has been set up by the project to test this approach. It gave an opportunity to create the M&E to evaluate the impact the project made and come up with certain policy recommendations (which can be presented in full as PPT presentation).

The overall objectives of the project were to address poverty, improve food security at household level and enhance environmental security through the development, promotion and adoption of strategies that enhance the productivity of existing irrigated farming systems in Central Asia. A subsequent objective was to test and introduce mechanisms for knowledge sharing and outscaling of project ‘bright spots’ practices and experiment results, such as LAs.

<table>
<thead>
<tr>
<th>Themes 2: Impact evaluation: Approaches and methods</th>
<th>Session number: [40.1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks</td>
<td></td>
</tr>
<tr>
<td>Session stream: Water supply and sanitation</td>
<td>Session name: Water supply and sanitation</td>
</tr>
</tbody>
</table>

**The impact of Water Supply and Sanitation interventions on child health in Nepal: evidence from the 2006 Nepal Demographic and Health Survey NDHS survey**

**Presenter: Ron Bose**

In this paper I examine the impacts on child health, using diarrhea as the health outcome, (amongst children living in households) with access to different types of water and sanitation facilities, and from socio-economic and child specific factors. Using cross-sectional health 2006 DHS survey data for Nepal, I employ statistical techniques to match children belonging to different ”treatment” groups, defined by water types and sanitation facilities, with children in a
"control" group. I also employ a variety of regression techniques to check for the robustness of my results, and demonstrate that it is important to incorporate the survey weights in both linear regression and matching estimation. Ignoring the survey design can affect estimates of population-level effects substantially in the analysis. The paper will provide guidance on ways in which matching can be used to estimate average cost-effectiveness of interventions. Finally I discuss some practical issues related to implementing the analysis using popular statistical software tools available in such packages as STATA.

| Themes 2: Impact evaluation: Approaches and methods | Session number: [40.1] |
| Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks | |
| Session stream: Water supply and sanitation | Session name: Water supply and sanitation 1 |


**Presenter: Howard White**

What is the relative effectiveness of water supply and sanitation interventions? What do we know about what works, what doesn’t and why? This presentation draws on two studies which address this question; the first a sectoral impact review conducted for the Independent Evaluation Group of the World Bank, and the second a 3ie-update of formal meta-analysis of impact studies in this sector. The findings point clearly to the benefits of sanitation interventions, and the relative lack of efficacy of community-based water supply as a health intervention. Point of use water treatment and hand washing, both cheap compared to community water supply, are found to be particularly effective.

**Water and Sanitation Interventions for better child health: evidence from a synthetic review**

**Presenter: Hugh Waddington**

This paper provides the results of a synthetic review of the effectiveness of interventions in water, sanitation and hygiene (WSH) in promoting better health outcomes in developing countries as measured by the incidence of diarrhoea among children. One billion people worldwide lack access to clean water, and 2.5 billion do without adequate sanitation. Interventions to effect improvements in WSH are therefore an important focus of efforts to improve quality of life in developing countries. By improving access to clean water and sanitation facilities and promoting better water-use, sanitation and hygiene practices, WSH interventions contribute towards better health outcomes by reducing risk of diarrhoeal disease, which is estimated to be the biggest killer of children globally.
There is a large and growing literature examining the effects of WSH interventions on outcomes in developing countries, utilising a range of study methodologies. WSH interventions are diverse, ranging from interventions to improve access to water through household and community provision; quality of water at source or point-of-use; sanitation facilities to promote safe disposal of waste; and interventions promoting safe hygiene practices. This review provides evidence on relative effectiveness of these interventions, by synthesising information from impact studies conducted during the last 30 years. It has been conducted to Campbell Collaboration (C2) methodological standard, and, in addition, draws on behavioural and contextual information in an attempt to explain the reasons for effectiveness and (likely) sustainability.

Themes 2: Impact evaluation: Approaches and methods
Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks

Session number: [40.1]  
Session stream: Water supply and sanitation  
Session name: Water supply and sanitation 1

Trickle down: Diffusion of chlorine for drinking water treatment in Kenya

Presenter: Clair Null

In this presentation we will emphasize the importance of collecting data on willingness to pay and the costs of the intervention as part of any rigorous evaluation of technology adoption. Credibly establishing that new technologies are effective is an important first step of impact evaluations, but policy makers need to know whether such technologies are valued by the populations they are intended to serve. Ultimately, data on the costs of providing a new technology and potential users’ willingness to pay for the new technology will determine whether or not adoption is sustainable and can help inform subsequent iterations of product design and/or pricing policies. We will discuss these issues in the context of a case study which identified a very effective health technology for which households nonetheless had a very low willingness to pay.

We study the distribution of dilute sodium hypochlorite to treat drinking water in homes in rural Kenya, and estimate how much households value clean water and how adoption is mediated by social networks. Using a randomized evaluation, we find large impacts of receiving a six-month supply of free chlorine on home water quality and child diarrhea rates. At the time of follow-up, 58% of treatment households had detectable chlorine in home water on an unannounced visit and there was a 69% reduction in water contamination as measured by the fecal indicator bacteria E. coli. Child diarrhea rates fell by 35-40%. However, data from a subset of households given coupons for discounted chlorine in local shops indicate that demand for the product is very low at even a minimal cost. We do not find that households with young children, who stand to benefit most from cleaner water, have a higher valuation for it. This discrepancy between the effectiveness and the willingness to pay for water quality improvements has important implications for policy makers when setting funding priorities. In particular, it seems unlikely that the significant benefits from point-of-use chlorination use can be realized in a market environment in which consumers have such low valuation of the good, and ongoing subsidies will likely be necessary to achieve sustained widespread adoption.

Nonetheless, even if only a few households adopt the product at the market price, with enough social learning widespread take-up might be achievable. Encouragingly, study households were more likely to talk with other community members about the product after the intervention, but ultimately we do not find strong evidence that these conversations translated into significantly higher take-up rates. Consistent with the low demand observed at positive prices, we find only a
moderate amount of social spillovers; these seem to have affected both treatment and comparison households. Though our social network effects are not very precisely estimated, close connections to members of the treatment group seem to have increased the probability that a household would adopt the WaterGuard technology. These network effects pale in comparison to the effect of assignment to treatment, but among comparison households the increase is economically relevant. While we are generally pessimistic about the potential for network spillovers to meaningfully increase take-up in the absence of a broader intervention, the finding that community leaders are particularly influential does identify a promising distribution channel in this area that could be more cost-effective than less targeted forms of social marketing.

<table>
<thead>
<tr>
<th>Themes 2: Impact evaluation: Approaches and methods</th>
<th>Session number: [40.2 ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks</td>
<td></td>
</tr>
<tr>
<td>Session stream: Water supply and sanitation</td>
<td>Session name: Water supply and sanitation 2</td>
</tr>
</tbody>
</table>

**Sound Impact Evaluations of German DC Programs: Experiences with Quasi-Experimental Designs and Mixed Methods Approaches. Case Studies from the Zambian and Kenyan Water Sector**

**Presenter: Nicola Reade**

In the past, German DC evaluations did not comply with increased standards discussed under the key word “rigorous impact evaluations” in the international DC evaluation community. Rigor and credible evidence on impacts were hardly ever provided due to lack of challenging methodological evaluation designs.

To meet these newly discussed demands and to have more sound facts and figures on achieved impacts, GTZ launched three pilot impact evaluations in the water sector in 2008, covering the countries Zambia, Kenya and Turkey. The impact evaluations were carried out within the standard framework of “External Independent Evaluations” commissioned by GTZ Head Office Evaluation Unit and conducted by the independent research institute Center for Evaluation.

The impact evaluations of the Water Sector Reform Programs in Zambia and Kenya, which will be discussed in the proposed presentation, were designed as interim program evaluations. The evaluation in Kenya was furthermore implemented as joint evaluation of the Kenyan Water Sector together with the Ministry of Water and Irrigation (MWI), Sida and Danida.

Both evaluations focused not only on measuring impacts but on measuring the general development policy effectiveness based on the DAC criteria relevance, effectiveness, impact, efficiency and sustainability. A quasi-experimental research design and a sound mixed methods approach was applied. The evaluation methodology comprised of: (a) Hypothesis based cause-effect analysis; (b) Quasi-experimental research design differentiating between target and comparison group; (c) Multi-method approach (document and file analysis, collection of statistical data, group discussions and guided interviews, standardized survey with members of the target and comparison group, field visits and observation); (d) Quantitative as well as qualitative data analysis methods.

To ensure quality and transparency of the evaluation process, the evaluation approach included adherence to DAC standards and a participative approach. According to the principles of research-triangulation, the evaluation questions were addressed by both international and national evaluators. Two teams conducted the evaluations under the lead of the national
expert from CEval: One team focusing on the overall data collection and reporting process with a duration of up to 80 days. The second team comprised of two experts and 10 enumerators focusing on conducting a standardized survey with approx. 1,000 target group members within a time period of max. 10 days.

The proposed presentation will focus on comparing and contrasting the two case studies, their planned and realized evaluation approaches and methods, as well as the constraints faced. Special emphasis will be put on assessing the strengths and weaknesses, possibilities and limitations of the approach and methods used and their influence on credibility of evidence and adherence to quality standards in impact evaluations.

Themes 2: Impact evaluation: Approaches and methods
Themes 6: Impact evaluations of policies, programs, projects, partnerships and networks
Session stream: Water supply and sanitation
Session name: Water supply and sanitation 2

Stories in Numbers: Using Financial Indicators to Reveal Household and Community Narratives in Impact Evaluation of Water Projects

Presenter: Michael P Canares

Impact evaluations are designed to determine the degree by which problems and needs faced by beneficiaries or clients are met by specifically designed interventions. They are also conducted to ascertain the degree of success of planned interventions, whether or not these are carried out as planned. In other cases, they are done to identify success of project interventions by looking into the desired impact on beneficiaries. Whatever be the reason for conducting the evaluation, such is done with the primary objective to help improve the current implementation or learn insights from its experience for application in future endeavors.

However, impact evaluations infrequently use financial indicators to gauge achievements of programs or projects in these areas. This paper argues, using the experience in evaluating water projects in the Philippines, that financial indicators are very useful tools not only in indicating financial and quantitative achievements, but also in revealing significant stories in the lives of water beneficiaries and the water user associations.

Financial indicators can be used to assess relevance, effectiveness, efficiency, impact and sustainability. Financial indicators like percentage of water costs to household weekly budget, water acquisition cost per unit, and labour efficiency ratios at the level of households and beneficiaries measure relevance, effectiveness, and efficiency at the micro-level. At the water-service provider level, financial indicators as water generation costs, cost per beneficiary, asset recovery ratio and turnover, replacement capacity, collection efficiency, and self-sufficiency ratios indicate not only the effectiveness and efficiency of water service providers but also their capacity to sustain service operations after assistance by funding organizations.

Using the financial perspective as a lens, this paper argues that impact evaluation may use financial indicators to make a judgement as to the congruence between needs and intervention, between assertions and actual situation, between current condition and future prospects. While not a substitute to other qualitative and quantitative measurements, financial indicators have the capacity of revealing important stories from the perspective of beneficiaries and communities and reinforce, if not enrich, our understanding of program or project inputs, process, achievements, issues, and concerns.
The Importance of a methodologically diverse approach to impact evaluation

**Presenter: Burt Perrin**

In December 2007, the European Evaluation Society issued a statement: *The Importance of a methodologically diverse approach to impact evaluation – specifically with respect to development aid and development interventions*. This presentation will provide the background to this statement, as well as discuss some of the issues that it raises.

For example, the statement indicates that the EES "supports multi-method approaches to impact evaluation and does not consider any single method such as RCTs as first choice or as the 'gold standard'.” This presentation, drawing upon the statement, will discuss in particular the importance of taking a multi-method approach, matching the method to the situation. It will consider some situations where an RCT may be appropriate, as well as limitations that may limit its applicability and ability to contribute to policy. As well, the EES statement observes that RCTs are based upon one particular model of causality, and that there are other well-established scientific models of causality that may be more appropriate in the case of complex interventions.

Finally, this presentation, drawing upon the EES statement, will note that in the context of the Paris Declaration, mandating a specific approach could undermine objectives of capacity building and in turn limit buy-in and support for evaluation and for subsequent action. In contrast, a multi-method approach that can draw from the rich diversity of existing frameworks and that engages the developing world can result in means of impact evaluation that are most likely to aid in the improvement of development effectiveness.

**Qualitative methods and impact evaluation**

**Presenter: Martin Prowse**

Calls for better aid effectiveness and for more evidence-based social policy have led to Randomised Control Trials (RCTs) and rigorous Impact Evaluations (IE) reaching the top of the international policy agenda. Despite sharing some characteristics, these two evaluation approaches differ in important respects. One of these relates to the role qualitative research could play in such evaluation practices. For example, the influential academics associated with the Poverty Action Lab at MIT appear to discount qualitative methods entirely. On the other hand, realist researchers associated with NONIE and 3IE initiatives are clear that qualitative methods should have a complementary, if secondary role to rigorous quantitative results. Until now, though, very little attention has been paid to the possibility of Impact Evaluations that allow a primary role to qualitative methods. This paper outlines a range of possible
qualitative methods that could be the primary tool within the rubric of a randomized control trial, suggests rigorous ways such data could be analysed and interpreted, and discusses the implications of such a shift for policy and programming.

Themes 1: Improving development effectiveness: the role of impact evaluation

Themes 2: Impact evaluation: Approaches and methods

Session number: [41.1]

Session stream: When and how to do impact evaluation

Session name: When and how to do impact evaluation

Rigorous Impact Evaluation –What It Is About and how it can be done in practice

Presenter: Alexandra Caspari

The term “impact evaluation” has become enormous popular. Especially the new add on “rigorous” provoked many methodological discussions: A lot of national and international, smaller and bigger conferences were hold and much more papers highlighting this topic were written. The discussion about rigorous impact evaluation got more and more specialized and sophisticated – for some maybe a little bit too complex: Most of those dealing with evaluations are neither exclusively focused on evaluating impact nor on methodological issues. But the underlying question or unresolved problem – How do we know when development is truly successful? – is much too important to limit the discussion to a few experts. Thus it is necessary to break down this topic into its simplest components to help better understand the key elements; to simplifying the problem into a form that can be readily approached by anyone interested in.

The presentation will start with a glance at the historical background of the discussion – the finding of the CGD publication “When Will We Ever Learn? Improving Lives Through Impact Evaluation” – and the consequential need to discuss how impact evaluations could be done better from a methodological point of view. What is meant by “clear causation”, “causal attribution”, and the term “counterfactual” will be explained and the necessity for adequate research designs will be shown.

The main focus of the presentation will be on commonly used and feasible impact evaluation designs: The differences between non-experimental, quasi-experimental, and experimental designs will be explained and their options and limits will be illustrated. It will be shown that it is not reasonable to prematurely reject quasi-experimental designs or RCTs as unnecessarily sophisticated or because of ethical concerns: The various realistic ways in which appropriate impact evaluation designs can be introduced in an ethically and politically acceptable manner in an ordinary impact evaluation will be described in detail: Matching on observables, the pipeline approach, regression discontinuity and even the underlying idea of propensity score matching will be explained by examples. The aim is to demonstrate that adequate designs (except for PSM) can be implemented in almost any impact evaluation by a professional evaluator and thereby enormously increase the quality of the findings.

Finally, the discussion about this specific methodological issue needs to be set in the context of an evaluation as a whole: An adequate evaluation design for a beneficiary survey does not give any answer to the question of why an intervention shows impact, the “black box” still remains. Thus, causal attribution is necessary but in no case sufficient. There are a lot of other methods that has to be used as in every evaluation. The point is: The discussion about rigorous impact evaluations by no means denies the necessity of the use of mixed methods.
Quand l’évaluation d’un projet pilote devient l’étude de faisabilité pour la reproduction du projet

*Presenter: Par Linda Hanna, Marie Gervais*

Le conflit du Darfour a déplacé dans l’Est du Tchad environ 200 000 réfugiés, et ceci dans un milieu naturel très fragile et quasiment dépourvu de ressources naturelles (eau et bois). L’afflux massif des réfugiés a exercé une pression supplémentaire sur le milieu : le partage des maigres ressources naturelles a entraîné de nombreux conflits entre populations locales et réfugiés. Dans ce contexte d’urgence, il était nécessaire de prendre des décisions rapides et de proposer une solution d’appui ou de remplacement à l’énergie traditionnelle constituée par le bois.

L’implantation du projet de cuiseurs solaires a donc été décidée sans étude préalable de faisabilité. Suite à cette expérience, il a paru nécessaire d’évaluer les impacts de cette initiative amorcée en contexte d’urgence.

La présente évaluation des expériences concrètes vient donc compenser l’absence d’analyse théorique préalable et pourra servir d’étude de faisabilité pour la mise en œuvre de projets similaires dans d’autres contextes.

Cette présentation retracera les grandes lignes d’une démarche d’évaluation de type formative, participative et qualitative réalisée dans le contexte d’un camp de réfugiés soudanais à l’est du Tchad. Les défis de planification et de mise en œuvre d’une évaluation dans un contexte en grande mouvance et présentant des ressources limitées seront explorés tout comme les stratégies de mobilisation et d’adaptation qui furent nécessaires pour assurer le succès de la démarche évaluative et l’amélioration rapide du projet-pilote. Les apprentissages réalisés à la fois pour l’équipe en place et pour l’évalutrice seront également présentées. À titre indicatif, voici quelques constats qui émergent de l’évaluation :

a) Le constat d’une situation environnementale problématique qui entraîne des impacts nuisibles et des violences physiques entre les communautés locales et les réfugiés ;

b) Le diagnostic social et environnemental réalisé en raccourci compte tenu de l’urgence d’agir pour prévenir la dégradation de l’environnement physique et social ;

c) Les premières démonstrations, formations et distributions des cuiseurs solaires, puis la diffusion de cette innovation à l’échelle du camp d’Iridimi ;

d) Les premiers constats en cours de réalisation et des modifications visant à améliorer le matériel et son utilisation : démarche exploratoire par essais-erreurs ;

e) L’évaluation du projet, puis la prise en compte des résultats de cette évaluation pour modifier certaines façons de faire et optimiser l’exportation future du projet.
Rethinking, Reshaping and Reforming Impact Evaluation for Development Effectiveness in the 21st Century

**Presenters:** Sanjeev Khagram, Zenda, Patricia Rogers, Alex Jacobs

Development and development effectiveness are dramatically different in the 21st century. They increasingly involve, among others:

- a broader range of views on the appropriate goals of development
- active engagement by a much wider range of actors and partners across issues, sectors and levels
- citizens as primary constituents and agents of development
- interventions which are not standardised or static but dynamic and adaptive
- plural development initiatives with a greater variety of intended impacts such as average income, sustainable livelihoods, gender equity, transparency and security, participation, and others
- a shift away from aid-based development
- recognition of the importance of global policies for global development.

These changes require us to **RETHINK** – How are development and development effectiveness changing? How should impact evaluation evolve in order to respond to these changes? **RESHAPE** - How should we conduct impact evaluation differently? What methods, tools and techniques need to be created, developed or more widely used? **REFORM** - What are the structural and systemic changes needed to support reshaped evaluation practice? How do we encourage these changes?

The session will explore how these changes have affected the way we understand development and development effectiveness, and the implications this has for development impact evaluation. Specific conversations around appropriate approaches, methods and tools for impact evaluation will follow the general overview. Participants will be introduced to a range of rigorous and credible methods that are likely to be particularly appropriate for development and development effectiveness in the 21st century, with the goal of making such methods better understood and more readily available. The session will conclude with some proposals for the future.

**Session design:** Panel discussion. Participants will be introduced to a series of propositions around:

- Changes in development and development effectiveness in the 21st century
- The implications this has for impact evaluation
- Appropriate approaches, methods and tools
- Systemic changes needed to impact evaluation to respond to these changes

**Themes:**

<table>
<thead>
<tr>
<th>Session stream: M&amp;E training in Tibet</th>
<th>Session name: M&amp;E training in Tibet</th>
</tr>
</thead>
</table>

**M&E Training in Tibet: Adapting to and Learning from Other Cultures**

**Presenter:** Laura P. Luo

Evaluation is an emerging, and fast growing field in China. Most international agencies that provide development aid to China require that evaluation be conducted for the projects to see the impact of the intervention. In recent years, the Chinese government has also called for establishing an open, accountable and transparent government. Evaluation, including impact evaluation of development projects, has become increasingly important.
The Basic Human Needs Project (BHNP) in Tibet was a bilateral poverty alleviation project funded by the Canadian International Development Agency (CIDA) and the Chinese government. BHNP covered farming, livestock rearing, rural health, water and sanitation, forestry, grassland management, community and institutional capacity building, with gender awareness and environmental protection as the cross-cutting themes for the project. The project also included a component to build evaluation capacity for officials in Poverty Alleviation and Development Office and its partner organizations, including line government agencies, in Tibet. This presentation will discuss the relevance of the training to impact evaluation of development projects required by the Chinese government. Additionally, the presentation will introduce to the audience the diverse, participatory methods used in the training to help build evaluation capacity of local government officials in Tibet. The presentation will also discuss the importance of learning about and tailoring training methods to local culture when conducting evaluation training.

This presentation will explore using a unique format, i.e., video, which will provide a glimpse of the beauty and richness of the Tibetan culture, poverty alleviation projects conducted in Tibet, the evaluation training and so on. The presenters hope that this innovative format will better help the audience learn about impact evaluation in a different culture.

<table>
<thead>
<tr>
<th>AfrEA Sessions</th>
<th>Session number: [44.1]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Session stream: Scholarly Journal focused on African evaluation practice</strong></td>
<td></td>
</tr>
</tbody>
</table>

There have been numerous calls by national associations and evaluation practitioners in Africa for AfrEA to establish an African evaluation journal. Africa has a vibrant and growing evaluation community, with very significant international interest and participation. A journal could provide an important voice for this community, and is in line with the mandate obtained at the January 2007 Fourth AfrEA Conference to “make evaluation Africa’s own”. AfrEA has proposed the establishment of an African evaluation journal, in an effort to meet the growing demand for systematic knowledge about the theory and practice of evaluation in developing countries.

AfrEA established a Journal Task Force that has developed a concept note detailing the vision and purpose, scope and focus, and management of “The Journal of African Evaluation”.

The Journal Task Force proposes a panel discussion at the Cairo Conference. Through this session the Task Force hopes to provide opportunity for the African evaluation community to participate in the formative processes of the journal. The Task Force aims to open for discussion some of the challenging issues regarding content, language, focus, funding and editorial processes, which need to be settled in the process of developing the journal. It also hopes to build interest, participation and support for the development processes that need to be undertaken.

**Methodology:** The following four presentations will be made, followed by a discussion:

**Making Evaluation Our Own Through Scholarly and Development-Oriented Evaluation**  
*Persenter: Sulley Gariba*

**What Kind of Content Might We Expect From An African Evaluation Journal?**  
*Persenter: Florence Etta*

**What It Takes to Start and Nurture A New Journal**  
*Persenter: Donna Mertens*

**Thinking Evaluatively About An African Evaluation Journal**  
*Persenter: Bob Williams*
The evaluation of development impacts is a science that is underpinned by many roots, including universally accepted norms and standards, as well as culturally-specific systems of knowledge, values and reflections of change and transformation. In the construction of systems of evaluation, Africa and African evaluators are challenged by the diversity of cultures and the evolution of developmental models that transcend time and traditions. This makes the challenge of ‘making evaluation our own’ both an imperative for the survival of evaluative thinking and practice in Africa; as well as Africa’s contribution to the richness of evaluation globally.

A further challenge lies in praxis – developing and utilizing conceptual understanding evaluations and African institutions to strengthen practice. This would require a strong foundation of evaluators with sufficient experience and knowledge and able to work together to influence and shape evaluation practice on the continent.

i) A Panel Discussion involving several expert panelists with experience in promoting and applying evaluation practice that is grounded in culturally specific systems of knowledge.

ii) Working Groups to draw on experiences and lessons to identify concrete actions and developing a Programme of work for ‘making evaluation our own.

Output: A broad framework, identifying concrete actions for ‘making evaluation our own’ in Africa, and influencing evaluation practice globally.