Monitoring and Evaluation
Topic Guide

GSDRC

Governance and Social
Development Resource Centre

Monitoring and Evaluation
Topic Guide

GSD RC
**Governance and Social
Development** Resource Centre

About the GSDRC

The Governance and Social Development Resource Centre (GSDRC) was established by the UK Department for International Development (DFID) in 2005 to provide access to high quality, timely information to support international development project and programme planning, policymaking, and other activities in the field
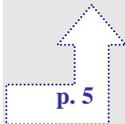
How to use this guide

This topic guide provides an overview of current knowledge in monitoring and evaluation of development activities. It includes short summaries of key texts. Each short summary links to an extended version in the appendix, along with information on how to access the original text. Both the short and extended summaries are cross-referenced in the guide.
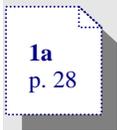
Key
1a   p. 28 This symbol, next to each short summary in the main text, indicates where the extended summary can be found in the appendix
This symbol, next to each extended summary in the appendix, links back to the short summary in the main text   **p. 5**
This symbol, at the end of the extended summaries in the appendix, indicates how to access the original text

This topic guide can also be accessed on the GSDRC website: www.gsdrc.org.

The online version includes guidance for DFID staff, evaluations of DFID's aid instruments, and evaluations of cross-cutting issues.

Contact: enquiries@gsdrc.org

Contents

www.gsdrc.org

1. Introduction

How can the impact of development programmes be assessed with a view to improving their efficiency and effectiveness? What particular challenges are involved in monitoring and evaluating development interventions, and how can these be addressed?

Monitoring and evaluation is vital to ensuring that lessons are learned in terms of what works and what doesn't in development. Understanding how and why programmes lead to desired (or undesired) outcomes can provide an evidence base for future policymaking and project planning. M&E also supports the need for aid agencies to be accountable for and demonstrate the results of their programmes.

Nevertheless, monitoring and evaluation has historically suffered from underinvestment, weak commitment to evidence-based policymaking, lack of incentives to carry out evaluations and a relative shortage of professional expertise. Even when methodologically sound evaluations are conducted, they are often under-utilized. Significant criticism has subsequently been levelled against the development community for failing to adopt a rigorous approach to understanding, and in turn being accountable for, the impact of their interventions.

This guide introduces some of the core debates and considerations for development practitioners involved in designing and managing monitoring and evaluation (M&E) activities. It provides case studies of applying different methodological approaches, tools for step by step planning, and lessons learned from international experience of M&E in a range of developing country contexts.

1.1. The Changing Context of M&E

The international context for M&E is changing, with increasing focus on measuring results and critically analysing aid effectiveness (see the Managing for Development Results initiative). Several international initiatives and agreements over the past decade have implications for development evaluation:

- The HIPC and PRSP Initiatives intend to integrate M&E better into funding decisions, and establish national M&E systems aligned to Poverty Reduction Strategies.

- The Millennium Development Goals now form the basis of progress indicators at the national level, and have meant much greater interagency cooperation around data collection.

- The Monterrey Declaration (2002) is often seen as the major impulse towards results-based monitoring and evaluation and the promotion of the allocation of aid based on the development of measures of effectiveness and results.

- The Paris Declaration (2004) committed OECD countries to increasing the country ownership of programmes, encouraging donors to harmonize and align with country monitoring and results frameworks.

Whilst conventional M&E has typically been dominated by donors, conducted at the project level without meaningful participation from government or civil society, the above initiatives have ushered in notable new trends:

- From monitoring and evaluating project processes, inputs and outputs to an emphasis on measuring results, outcomes and impact.

- From monitoring and evaluating projects, to new emphasis on evaluating the combined effects of aid - in part prompted by the increase in multi-donor programmes and sector-wide approaches.

- From M&E being predominately donor-led to increased interest in country-led approaches, with evaluations increasingly conducted in partnership with a broader range of stakeholders, including the programmes' intended beneficiaries.

- A renewed emphasis on sharing and using the results of evaluations effectively.

Crucially, however, new aid modalities such as budget support and country-led approaches present potential contradictions for the development community, requiring them to reconcile their need for accountability and a more rigorous focus on results with a "hands-off" approach to aid.

1.2. Common Tools and Approaches

Monitoring and Evaluation can take place at the project, programme, sector, or policy level. It can be undertaken by programme managers or external consultants and will usually involve local stakeholders. It is broadly considered good practice to develop an M&E Framework (the strategy for planning and organising the vital M&E activities) during the design phase of projects or programmes.

Most projects or programmes now formulate a 'logical framework' (or results framework), which hypothesises the causal link between a project/programmes inputs and outcomes. Nevertheless, there is no universal approach to designing or implementing M&E, and a wide range of tools are used in practice.

World Bank Operations Evaluation Department, 2002, 'Monitoring & Evaluation: Some Tools, Methods and Approaches', The World Bank, Washington, D.C

1a
p. 28

Monitoring and Evaluation (M&E) is an area of growing importance for the development community. It allows those involved in development activities to learn from experience, to achieve better results and to be more accountable. This report from the World Bank Operations Evaluation Department provides an overview of some of the M&E tools, methods and approaches on offer to development practitioners.

1.3. Terms and Definitions

Aid agencies use their own distinct terminology to describe their M&E systems and processes. Nevertheless, the OECD DACs definitions of key terms are widely accepted:

> Organisation for Economic Co-operation and Development –
> Development Assistance Committee (OECD-DAC), 2002, 'Glossary of
> Key Terms in Evaluation and Results-based Management', Working Party
> on Aid Evaluation, OECD-DAC, Paris
>
> The DAC Working Party on Aid Evaluation (WP-EV) has developed this
> glossary of key terms in evaluation and results-based management because of the
> need to clarify concepts and to reduce the terminological confusion frequently
> encountered in these areas. With this publication, the WP-EV hopes to facilitate
> and improve dialogue and understanding among all those who are involved in
> development activities and their evaluation, whether in partner countries,
> development agencies and banks, or non-governmental organisations. It should
> serve as a valuable reference guide in evaluation training and in practical
> development work.
>
> 📄 Full text available online: http://www.oecd.org/dataoecd/29/21/2754804.pdf

Monitoring is the ongoing, systematic collection of information to assess progress
towards the achievement of objectives, outcomes and impacts. It can signal potential
weaknesses in programme design, allowing adjustments to be made. It is vital for
checking any changes (positive or negative) to the target group that may be resulting
from programme activities. It is usually an internal management activity conducted by the
implementing agency.

Evaluation is defined by the OECD-DAC as:
"The systematic and objective assessment of an on-going or completed project,
programme or policy, its design, implementation and results. The aim is to determine the
relevance and fulfilment of objectives, development efficiency, effectiveness, impact and
sustainability." (OECD-DAC, 2002)

Evaluations tend to look at project impact and sustainability. They can be
periodic/formative - conducted to review progress, predict a project's likely impact and
highlight any necessary adjustments in project design; or terminal/summative -carried out
at the end of a project to assess project performance and overall impacts. They can be
conducted internally, externally by independent consultants (particularly in the case of
impact evaluations), or as a joint internal/external partnership.

1.4. Where is a good place to start?

For an introduction to both the context and application of M&E in development, see:

> Kusek, J., and Rist, R., 2004, 'Ten Steps to a Results-based Monitoring
> and Evaluation System', World Bank, Washington, D.C.
>
> Governments and organisations face increasing internal and external pressures to
> demonstrate accountability, transparency and results. Results-based monitoring
> and evaluation (M&E) systems are a powerful public management tool to achieve
> these objectives. This handbook from the World Bank presents a ten-step
> model that provides extensive detail on building, maintaining and sustaining a
> results-based M&E system.

6

2. Monitoring poverty reduction strategies

Progress towards national poverty reduction goals set out in Poverty Reduction Strategies (PRS) is monitored and assessed by governments, donors and civil society to determine the effectiveness of government expenditure, guide aid allocations, or inform and stimulate national debate on development progress. However, PRS monitoring has historically suffered from underinvestment, lack of co-ordination and weak national capacity.

There is considerable debate about how to overcome these constraints, which institutional arrangements for monitoring work best, and how different actors can contribute to the process.

2.1 Poverty Monitoring Systems

The institutions, relationships and systems established to monitor poverty reduction are often referred to as the 'Poverty Monitoring System'. Underinvestment in these systems has arguably resulted in a major shortfall in understanding the effectiveness of poverty reduction efforts.

Holvoet, N., and Robrecht, R., 2006, 'Putting the New Aid Paradigm to Work: Challenges for Monitoring and Evaluation', Institute of Development Policy and Management, Antwerp

2a
p. 31

What are the challenges in monitoring and evaluating Poverty Reduction Strategy Paper (PRSP) outcomes? Have issues been neglected in these key areas? This study by the Institute of Development Policy and Management, University of Antwerp, reviews and assesses M&E systems for 11 Sub-Saharan African countries. These systems are embryonic in the countries reviewed, and are a weak link in most PRSPs. The emphasis on participation and comprehensiveness places unrealistic demands on national capacity, exacerbated by the scant attention given to M&E by donors.

2.2. Lessons Learned in Design

Setting up a poverty monitoring system means defining goals, indicators and targets, agreeing data requirements and deciding who will collect data and how. In reality, the different components that make up the monitoring system can have independent origins, and there is often little coherence across the system.

World Bank, 2002, 'Chapter 3: Monitoring and Evaluation', in PRSP Sourcebook, World Bank, Washington, D.C.

2b
p. 32

How do we know if a poverty reduction strategy is effective? First, a poverty monitoring system is needed to track key indicators over time and space and to determine if they change as a result of the strategy. Second, rigorous evaluations should be done selectively to assess the impact on poverty of interventions that are key components of the strategy. This chapter from the World Bank's Poverty Reduction Strategy Paper Sourcebook examines the features of poverty monitoring systems and explores some of the key issues which arise during implementation.

7

Lucas, H., Evans, D. and Pasteur, K., 2004, 'Research on the Current State of PRS Monitoring Systems', IDS Discussion Paper no. 382, Institute of Development Studies, Brighton

What are the key factors that determine the success or failure of Poverty Reduction Strategies (PRS) monitoring efforts? Who should be involved? How should the information gathered be put to use? This study by the Institute of Development Studies (IDS) reviews the main issues arising from the implementation of PRS monitoring systems. It draws on the experience of different countries and suggests possible monitoring approaches to suit specific contexts.

Booth, D. and Lucas, H., 2002, 'Good Practice in the Development of PRSP Indicators and Monitoring Systems', Overseas Development Institute, London

How can poverty reduction strategies best be monitored? And how can monitoring lead to greater success in reducing poverty? This study by the Overseas Development Institute (ODI) draws conclusions about best practice from a review of Poverty Reduction Strategy Papers (PRSPs) and suggests ways in which monitoring can be improved. PRSP monitoring calls for fresh thinking. It needs to be geared to what is new and challenging about the PRSP initiative – particularly the effort to engage a wider range of stakeholders in policy dialogue about poverty reduction at the national level. It also needs an understanding of the relevant policy processes and the possible uses of information in enforcing new kinds of accountability and learning about poverty reduction.

2.3. Indicators

Indicators are the quantifiable measures of progress towards the intended inputs, outputs, outcomes and impacts of a project, programme or strategy. They are the measures for assessing the quantitative and qualitative impact of development efforts.

Reaching collective agreement on indicators is a complex task, not least because positive and negative changes (in relation to project objectives and real world outcomes) have varying time frames of measurement/assessment, and developing credible short-term indicators of long-term change can be problematic.

Arguably the most important criteria for selecting indicators is that they are realistic - it is essential that the necessary time, resources and data are available to measure them. It is also widely acknowledged that indicators should be independent of possible bias of the observer, and fulfil the SMART criteria - 'Specific - Measurable - Achievable - Relevant – Timebound'.

> World Bank, 2004, 'Poverty Monitoring Guidance Note 1: Selecting Indicators', World Bank, Washington, D.C.
>
>
> What is the best method for selecting appropriate indicators to monitor the implementation of poverty reduction strategies? This paper from the World Bank provides a summary of good practice for selecting such indicators. In general, it is preferable to select a few good quality indicators which are easily measurable within the current capacity and which cover the right questions, at the right level of disaggregation. This is ultimately a political process which needs to take into account existing constraints.

Indicators can only be used effectively to promote development outcomes if they are fully understood by, and command widespread support among, a broad range of national stakeholders.

> Scott, C. and Wilde, A., 2006, 'Measuring Democratic Governance: A Framework for Selecting Pro-poor and Gender Sensitive Indicators', United Nations Development Programme, Oslo
>
>
> How should we measure democratic governance? Most indicators are developed by external stakeholders to compare nation states and are not designed to help countries undertake governance reforms. This UNDP guide presents a framework for generating pro-poor gender sensitive indicators to help policy makers monitor and evaluate democratic governance at the country level. It argues that indicator selection is itself a governance process.

For guidance on selecting and using indicators in specific sectors, see:

GSDRC Guide to Measuring Governance
http://www.gsdrc.org/go/topic-guides/indicators

GSDRC Guide to Rights Monitoring and Indicators
http://www.gsdrc.org/index.cfm?objectid=8940BB1B-CEDA-66E8- EF72730D3EA7A360

GSDRC Guide to Quantifying Social Exclusion
http://www.gsdrc.org/index.cfm?objectid=E097D531-AE55-969D-FDE7162AE1C21CB6

2.4. Institutions for PRS Monitoring

Poverty monitoring systems typically involve a broad range of actors, including data producers, analysts and users from government, donors, consulting firms, think tanks and civil society. In this environment, a clear, coherent system of roles, responsibilities and reporting mechanisms is necessary to avoid fragmentation and ensure that information flows effectively back in to policymaking.

> Bedi, T., et al., 2006, 'Beyond the Numbers: Understanding the Institutions for Monitoring Poverty Reduction Strategies', World Bank, Washington, D.C.
>
> Monitoring systems are central to the effective design and implementation of a Poverty Reduction Strategy (PRS). Many existing PRS monitoring systems lack coordination and a coherent institutional framework linking monitoring and decision making. This World Bank report draws on 12 country studies to conclude that PRS monitoring systems should build on existing elements to begin a process of gradual change. Clearly defined relationships, incentives and activities and identification of entry points in decision-making processes facilitate the supply of monitoring information and its integration into improving PRS policies.

2g
p. 40

Case Study: The Poverty Monitoring System in Uganda

The Uganda Poverty Monitoring System has seen some success in improving country ownership, increasing results orientation and promoting partnership.

> Booth, D., and Nsabagasani, X., 2005, 'Poverty Monitoring Systems: An Analysis of Institutional Arrangements in Uganda', Overseas Development Institute, London
>
> Monitoring activities can play an important role in promoting country ownership of poverty-reduction policies where they are closely related to a politically supported and maturing budget process. Focusing on Uganda, this paper from the Overseas Development Institute (ODI) is one of a series of country studies intended to inform the design and implementation of poverty monitoring systems (PMS). While Uganda had an early start in creating a political environment that enabled results to influence policy, problems with incentives and partnership continue. There will inevitably be setbacks in developing a PMS - a process that requires coordination, policy-relevance and strong donor-country partnerships.

2h
p. 41

2.5. What Role for Donors?

PRS monitoring has recieved criticism for being a process primarily set up to meet donor accountability requirements and support conditionality. Pressure for effective monitoring systems largely comes from donors, rather than from domestic stakeholders. There is considerable debate about what role donors can usefully play in supporting the monitoring process, in particular how they can adopt a country-led approach whilst meeting their own accountability requirements.

> Lucas, H., and Zabel, M., 2005, 'The Role of Donors in PRS Monitoring Systems: Advantages and Disadvantages of the Different Forms of Involvement', Report prepared for the World Bank, HLSP Institute, London
>
> How can donors improve the way in which they engage with the monitoring of Poverty Reduction Strategy (PRS) implementation? This draft report prepared for the World Bank by the HLSP Institute considers the reported activities and experiences of donors in PRS monitoring. It assesses the current position and provides some guidance as to possible ways forward. The challenge of PRS monitoring involves the need to (a) provide up to date, reliable information; (b)

2i
p. 43

10

encompass both PRS implementation processes and poverty outputs/outcomes; and (c) persuade donors of the quality of information. Whilst the desirability of country ownership and harmonisation and alignment is well established, there remains a substantial gap between theory and practice.

2.6. Civil Society Participation

Civil society has an important role to play in terms of overseeing and helping to hold governments to account for progress towards national poverty reduction goals. Nevertheless, civil society organisations often face significant obstacles in systematically monitoring poverty reduction strategies where there is a weak culture of accountability or restricted access to information. How civil society participation and co-ordination can be improved is a key concern for civil society and donors alike.

Alliance Sud, 2006, 'Civil Society's Perspective on their Involvement in PRSP Monitoring: Assessing Constraints and Potentials in Nicaragua', Swiss Alliance of Development Organisations (Alliance Sud), Bern

2j
p. 44

What are the constraints and potentials facing civil society participation in Poverty Reduction Strategy (PRS) monitoring processes in Nicaragua? This Alliance Sud study reports on an ongoing initiative to develop an analytical tool to assess civil society's participation in PRS monitoring processes from their own perspective. It finds that basic conditions for a significant role for civil society organisations (CSOs) in the monitoring system in Nicaragua are not met.

2.7. Strengthening National Statistical Capacity

Good statistics are vital for the effective monitoring of development programmes and strategies, and ultimately to support evidence-based policymaking. However, many developing countries lack the institutional capacity or effective systems for gathering data. Many existing statistical systems were predominately designed to meet immediate rather than long-term data needs, and therefore lack co-ordination.

There is increasing recognition of the need for a strategic approach to statistical capacity development, particularly following the Second International Roundtable on Managing for Development Results in February 2004 and the resulting Marrakech Action Plan for Statistics (MAPS). This emphasised the need to develop National Strategies for the Development of Statistics (NSDSs).

Paris21 Secretariat, 2004, 'A Guide to Designing a National Strategy for the Development of Statistics (NSDS)', Prepared by the Partnership in Statistics for Development in the Twenty-first Century (PARIS 21), Paris

2k
p. 46

There is increasing awareness of the need to strengthen statistical capacity to support the design, monitoring and evaluation of national development plans. National Strategies for the Development of Statistics (NSDS) are designed to achieve this goal. This guide prepared by the Partnership in Statistics for Development in the Twenty-first Century (PARIS21) aims primarily to assist developing countries to design their NSDSs but will also be useful to development partners.

3. Impact Evaluation

The recent emphasis on accountability and results-based management has stimulated interest in evaluating not just the process, outputs and outcomes of development programmes, but also their impact (ultimate effect) on people's lives. Impact evaluations go beyond documenting change to assess the effects of interventions on individual households, institutions, and the environment, relative to what would have happened without them – thereby establishing the counterfactual.

This rigorous approach to evaluation is increasingly advocated as the only reliable way to develop an evidence base of what works and what doesn't in development. Nevertheless, impact evaluations remain relatively rare, and in practice evaluation methodology often has to be adapted to technical, political and capacity constraints.

3.1. Attribution and the Counterfactual: The Case for More and Better Impact Evaluation

Development interventions are not conducted in a vacuum, and it is extremely difficult to determine the extent to which change (positive or negative) can be attributed to the intervention, rather than to external events (such as economic, demographic, or policy changes), or interventions by other agencies.

Impact evaluations attempt to attribute change to a specific programme or policy and establish what would have happened without the intervention (the counterfactual) by using scientific, sometimes experimental, methodologies such as randomized control trials or comparison groups.

The Evaluation Gap Working Group at the Centre for Global Development is the foremost advocate for the use of such rigorous evaluation methodologies to address what it views as a critical gap in evidence about the real impact of development programmes.

> Center for Global Development, 2006, 'When Will We Ever Learn? Improving Lives Through Impact Evaluation', Evaluation Gap Working Group, Centre for Global Development, New York
>
> **3a**
> p. 48
>
> Despite decades of investment in social development programmes, we still know relatively little about their net impact. So why are rigorous social development impact evaluations relatively rare? This paper from the Center for Global Development (CGD) aims to address this question and provide recommendations for more and better evidence for policymaking and programme planning. A new, collective approach is needed, in which developing country governments, bilateral and multilateral development agencies, foundations and NGOs work together to define an agenda of enduring questions and fund the design and implementation of rigorous impact evaluations in key sectors.

3.2. Randomized Control Trials: The Gold Standard?

Randomized Control Trials (RCTs) are often referred to as the 'gold standard' of impact evaluation, but whether or not they are always feasible, appropriate and rigorous is the subject of some debate.

> Duflo, E., and Kremer, M., 2003, 'Use of Randomization in the Evaluation of Development Effectiveness', Paper prepared for the World Bank Operations Evaluation Department (OED) Conference on Evaluation and Development Effectiveness 15-16 July, 2003, Massachusetts Institute of Technology , Cambridge, Mass
>
> Just as randomised pharmaceutical trials revolutionised medicine in the 20th Century, randomised evaluations could revolutionise social policy in the 21st. This paper, prepared for a 2003 World Bank Operations Evaluation Department (OED) conference, draws on evaluations of educational programmes. It argues that there is an imbalance in evaluation methodology and recommends greater use of randomised evaluations. As credible impact evaluations, these could offer valuable guidance to international organisations, governments, donors and NGOs in the search for successful programmes.

3b
p. 50

Case Study: Progresa in Mexico

The Progresa case is considered one of the most successful examples of the application of a randomized control trial in a development context.

> Attanasio, O., Meghir, C., and Santiago, A., 2005 'Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate Progresa', Institute for Fiscal Studies (IFS), London
>
> What impact have monetary incentives had on education choices in rural Mexico? How can the design of educational interventions aimed at improving educational participation be improved? This paper from the Institute for Fiscal Studies (IFS) analyses the education component of the Mexican governments welfare programme, Progresa, which aims to reduce rural poverty. It argues that increasing the grant for secondary school children while eliminating it at the primary age would strengthen Progresas impact.
>
> *For further case studies of randomized control trials, see:*
> *http://www.povertyactionlab.com*

3c
p. 51

3.3. Adapting to Time, Budget and Data Constraints

Ideological positions can obscure the issue of what methodologies are actually feasible. Scientific approaches can be costly, time consuming, and therefore unrealistic. Many organisations don't have the resources to carry out the ideal evaluation, and an M&E framework needs to be designed with organisational capacity, human and financial resources and political context in mind.

It is important to understand the minimum methodological requirements for evaluation rigour in cases where it is not possible to use strong evaluation designs.

Bamberger, M., 2006, 'Conducting Quality Impact Evaluations Under Budget, Time and Data Constraints', World Bank, Washington, D.C.

How do cost, time and data constraints affect the validity of evaluation approaches and conclusions ? What are acceptable compromises and what are the minimum methodological requirements for a study to be considered a quality impact evaluation? This booklet from the World Bank provides advice for conducting impact evaluations and selecting the most rigorous methods available within the constraints faced. It provides suggestions for reducing costs and increasing rigour and clarifies the nature of trade-offs between evaluation rigour and budgets, time and data.

Bamberger, M., Rugh, J., and Mabry, L., 2006, 'Bringing It All Together: Applying Realworld Evaluation Approaches to Each Stage of the Evaluation Process' Chapter 16 in Realworld Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage Publications, California

How can the principles of optimal evaluation design be applied under real world conditions with budget, time, data and political constraints? This paper, adapted from chapter 16 of RealWorld Evaluation: Working under Budget, Time, Data and Political Constraints provides an overview of the RealWorld Evaluation (RWE) approach. It addresses constraints through practical suggestions applicable to both developing and developed world research. Understanding the aims and purpose of the evaluation, as well as the local context, is critical.

Evaluation designs need to adapt to local realities, and experience demonstrates that no single methodology is applicable in all cases.

White, H., 2006, 'Impact Evaluation: The Experience of the Independent Evaluation Group of the World Bank', World Bank, Washington, D.C.

Aid spending is increasingly dependent on proof that interventions are contributing to the attainment of Millennium Development Goals (MDGs). Yet there is still debate over the definition of impact evaluation and how it should be carried out. This paper draws on the experience of the Independent Evaluation Group (IEG) of the World Bank. It defines impact evaluation as a 'counterfactual analysis of the impact of an intervention on final welfare outcomes' and recommends a theory-based approach. Two sources of bias are highlighted: contamination and self-selection bias.

3.4. Mixed-Method Designs

The debate on whether randomized control trials are the 'gold standard' of impact evaluation is largely based on the assumption that only quantitative evaluation designs are rigorous. In reality, effective evaluations may require both qualitative and quanititative analysis of results. Combining both types of analysis (using the 'mixed-method approach') can overcome the weaknesses inherent in either method.

14

Bamberger, M., Rugh, J., and Mabry, L., 2006, 'Mixed-Method Evaluation' Chapter 13 in Realworld Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage Publications, California

Quantitative and qualitative methods of research each have strengths and weaknesses when applied in isolation. However, combining the two approaches through mixed-method evaluation is gaining wider acceptance among social science researchers as a way of conducting more comprehensive and robust analysis. This chapter from RealWorld Evaluation: Working Under Budget, Time, Data and Political Constraints discusses the most appropriate contexts and strategies for using a mixed-method approach. It argues that mixed-method evaluation is a flexible and practical technique which can be used at any stage of an evaluation. Nevertheless, a fully integrated approach requires extensive planning and deliberation to ensure that the most appropriate combination of methods is chosen and successfully implemented.

3.5. Toolkits

Baker, J., 2000, 'Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners', World Bank, Washington, D.C.

There is broad evidence that developmental assistance benefits the poor, but how can we tell if specific projects are working? Have resources been spent effectively? What would have happened without intervention? This comprehensive handbook from Directions in Development seeks to provide tools for evaluating project impact. It advises that effective evaluations require financial and political support, early and careful planning, participation of stakeholders, a mix of methodologies and communication between team members.

The World Bank, 2006, 'Impact Evaluation and the Project Cycle', PREM Poverty Reduction Group, World Bank, Washington, D.C.

The goal of an impact evaluation (IE) is to attribute impacts to a project using a comparison group to measure what would have happened to the project beneficiaries had it not taken place. The process of identifying this group, collecting the required data and conducting the relevant analysis requires careful planning. This paper from the World Bank provides practical guidance on designing and executing IEs. It includes some illustrative costs and ideas for increasing government buy-in to the process.

4. Managing M&E

Badly designed and managed evaluations can do more harm than good - misleading results can undermine the effective channelling of resources for poverty reduction.

Establishing international standards for methodological rigour, ethical practice and efficient management processes in monitoring and evaluation is an ongoing challenge. Key concerns are how aid agencies should oversee evaluations outsourced to consultants, how to build country ownership of M&E processes where there are significant capacity constraints or limited buy-in, and how to co-ordinate evaluations of joint donor programmes effectively.

4.1. Steps in Planning and Design

Monitoring and evaluation activities are usually broken down into stages of planning, implementation, analysis, dissemination and use.

Kusek, J., and Rist, R., 2004, 'Ten Steps to a Results-based Monitoring and Evaluation System', World Bank, Washington, D.C.
Governments and organisations face increasing internal and external pressures to demonstrate accountability, transparency and results. Results-based monitoring and evaluation (M&E) systems are a powerful public management tool to achieve these objectives. This handbook from the World Bank presents a ten-step model that provides extensive detail on building, maintaining and sustaining a results-based M&E system.

4a
p. 63

Aid agencies largely work to their own internal requirements for reviewing, reporting on and evaluating the inputs, process and results of their activities, producing internal guidance notes that describe the practical steps involved.

Department for International Development, 2005, 'Guidance on Evaluation and Review for DFID Staff', Evaluation Department, DFID, London

Good evaluation practice depends on a solid partnership between those commissioning and managing evaluation studies and the consultants undertaking the work and producing reports. This guide from DFID aims to improve the quality of decentralised evaluation. It outlines the steps for designing, managing, reporting on and responding to an evaluation.

4b
p. 64

United Nations Development Programme Evaluation Office, 2002, 'Handbook on Monitoring and Evaluating for Results', UNDP, New York
Since 1999, the United Nations Development Programme (UNDP) has placed greater emphasis on results in its work to eliminate poverty. That shift has led to new demands on Monitoring and Evaluation (M&E) in country offices and programme units. This handbook outlines an M&E framework for use by UNDP staff and partners that promotes learning and performance measurement.

4c
p. 66

4.2. Ensuring Evaluation Quality

Ensuring the quality and integrity of evaluation design is vital for reaching accurate and reliable conclusions about what works and what doesn't work in development. International standards emphasise the need for impartiality, appropriately skilled experts conducting the evaluation, participation, country ownership and timeliness (evaluations should be appropriately timed to influence policymaking).

The OECD DAC Network on Development Evaluation's principles for evaluation of development assistance are widely cited.

Organisation for Economic Co-operation and Development – Development Assistance Committee (OECD-DAC), 1991, 'Principles for Evaluation of Development Assistance', OECD-DAC, Paris

4d
p. 67

Aid evaluation plays an essential role in efforts to enhance the quality of development cooperation. This paper from the OECD's Development Assistance Committee presents a set of principles on the most important requirements of the evaluation process. Development assistance is a cooperative partnership between donors and recipients. Both must take an interest in evaluation to improve the use of resources through learning and to ensure accountability to political authorities and the public.

There is also a need to ensure that evaluations are conducted ethically, in a culturally sensitive manner that protects the anonymity and confidentiality of individual informants.

United Nations Evaluation Group, 2005, 'Standards for Evaluation in the UN System', UNEG, New York

4e
p. 69

An effective evaluation process is an integral part of any project. But what are the key elements of a successful and sustainable evaluation approach? This document produced by the United Nations Evaluation Group (UNEG) offers solid guidelines for evaluation planning, design, implementation and reporting. Fundamental requirements include: institution-wide support, clearly-defined and transparent responsibilities, appropriately qualified staff, and a constant commitment to the harmonisation and updating of methods used.

4.3. Identifying Threats to the Validity of Evaluation Findings

Significant criticism has been levelled against the development community for failing to adopt methodologically sound approaches to evaluating their activities. These include weak analysis of qualitative data and not paying enough attention to mapping the causal chain from inputs to impacts.

White, H., 2005, 'Challenges in Evaluating Development Effectiveness', IDS Working Paper 242, Institute of Development Studies, Brighton

4f
p. 71

Evaluation has a crucial role to play in today's results-based culture and in the context of the Millennium Development Goals (MDGs). How then, can the quality of evaluation be improved? This working paper from the Institute of Development Studies (IDS) argues that there has been inadequate investment in

17

methodology, often resulting in low quality evaluation outputs. It discusses techniques in three areas of contemporary relevance: measuring agency performance; evaluation methods at the project level; and sustainability analysis.

The validity and usefulness of an evaluation are determined, amongst other things, by its statistical validity, use/action orientation, transferability and fittingness.

Bamberger, M., Rugh, J., and Mabry, L., 2006, 'Strengthening the Evaluation Design and the Validity of the Conclusions' Chapter 7 in Realworld Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage Publications, California

4g
p. 72

How can threats to the validity of evaluations be identified and addressed? This chapter from 'Realworld Evaluation: Working Under Budget, Time, Data and Political Constraints' outlines some of the most common threats to the validity of both quantitative (QUANT) and qualitative (QUAL) evaluation designs. It offers recommendations on how and when corrective measures can be taken to ensure validity.

See also: Bamberger, M., Rugh, J., and Mabry, L., 2006, 'Integrated Checklist for Assessing the Adequacy and Validity of Quantitative, Qualitative, and Mixed-Method Design', Appendix 1 in Realworld Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage Publications

Whilst many of these methodological issues are recognised, overcoming them is difficult in practice, especially where there are limited funds designated for evaluation.

Green, A., and Kohl, R., 2007, 'Challenges in Evaluating Democracy Assistance: Perspectives from the Donor Side', Democratization, vol 14, no. 1, pp. 151-165

4h
p. 73

Why is there a lack of credible research into the impact of democracy assistance? What are the obstacles to conducting such research? This article from the journal Democratization shares insights from a donor-sponsored workshop on the challenges facing the evaluation of democracy and governance (DG) programming and assistance. It argues that the lack of credible research is partly due to a fundamental difference in orientation between the retrospective approach of academics and the prospective approach of donor agencies.

4.3. Evaluating Multi-donor Programmes

The Paris Declaration commits donors to co-operation and harmonisation in all stages of the development cycle. Joint evaluations are necessary where multiple agencies are involved in a chain of interventions to pursue similar outcomes, or to understand the combined effects of all interventions across a particular sector.

Joint evaluations present opportunities for donors to pool their technical and financial resources into more rigorous, in depth and longer-term evaluations and in turn reduce the multiple information demands on governments and stakeholders. Nevertheless, they require reconciling the often divergent mandates and preferred evaluation approaches of different agencies.

18

Organisation for Economic Co-operation and Development – Development
Assistance Committee (OECD-DAC), 2006, 'Guidance for Managing Joint
Evaluations', DAC Evaluation Series, Paris

Joint evaluations have become central to development practice in recent years.
Collective assessment of agencies' combined work minimises transaction costs
for developing country partners and addresses the large aid-giving role of joined-
up modalities such as basket funds and joint assistance strategies. This booklet
produced by the Development Assistance Committee of the Organisation for
Economic Cooperation and Development (OECD-DAC) provides practical
guidance for making joint evaluations efficient, educational and collaborative.

Organisation for Economic Co-operation and Development – Development
Assistance Committee (OECD-DAC) Network on Development Evaluation,
2007, 'Sourcebook for Evaluating Global and Regional Partnership Programs:
Indicative Principles and Standards', OECD, Paris

Global and Regional Partnership Programmes (GRPPs) are an increasingly
important modality for channeling and delivering development assistance. This
World Bank Independent Evaluation Group (IEG) Sourcebook, prepared under
the auspices of the OECD/DAC Network on Development Evaluation, is
designed to address the growing need for consensus principles and standards for
evaluating GRPPs. It comprehensively presents, synthesises, applies and
elaborates on existing principles and standards, aiming to improve evaluation
independence and quality. As a result, GRPPs should become more relevant and
effective.

4.4. The Politics of Evaluation

Evaluations are more than a technical process. They have the capacity to determine
access to resources and the funding fate of programmes. It is inevitable therefore that they
will be subject to pressures from different stakeholders to produce favourable
assessments or to avoid addressing sensitive issues.

Bamberger, M., Rugh, J., and Mabry, L., 2006, 'Reconciling Different Priorities
and Perspectives: Addressing Political Influences', Chapter 6 in Realworld
Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage
Publications, California

No evaluation can ever be value free and completely objective. Decisions as to
what to study, which methods to use and whose criteria define programme
success, all involve human judgement. This chapter from 'RealWorld Evaluation:
Working Under Budget, Time, Data and Political Constraints', discusses how
political factors affect evaluation. It provides a detailed analysis of possible
pressures and constraints in evaluation design, implementation, dissemination
and use.

4.5. Promoting the Use of Evaluation Findings

M&E should ultimately result in improved policy and practice. Yet the findings and recommendations of evaluations are frequently underutilized. In order for evaluations to be influential, it is important to consider how to integrate them into the policymaking cycle, the political incentives to take up findings, and how the report is presented and understood by different stakeholders.

World Bank Operations Evaluation Department, 2004, 'Influential Evaluations: Evaluations that Improved Performance and Impacts of Development Programs', World Bank, Washington, D.C.

Evaluations can be a cost-effective way of improving the performance and impact of development activities. However, they must be conducted at the right time, focus on key issues and present results in an accessible format. This report from the World Bank Operations Evaluation Department presents eight examples of evaluations that have had an important impact, and summarises lessons learned.

4l
p. 80

The relative influence of evaluations also depends on whether there are sufficient institutional arrangements to support the transformation of policy lessons into policy actions.

Gordillo, G., and Andersson, K., 2004, ' From Policy Lessons to Policy Actions: Motivation to Take Evaluation Seriously', Public Administration and Development, Volume 24, pp. 304-320

Whilst recent political reforms have sometimes led to modifications in countries' national policies, the link between policy evaluation and policy actions is often weak. So why do so many governments take policy evaluation so lightly? This article from the Food and Agriculture Organisation (FAO) and the Center for the Study of Institutions, Population, and Environmental Change (CIPEC) at Indiana University analyses the institutional aspects of creating effective systems for monitoring and evaluation (M&E) in government-led rural development efforts in Bolivia and Brazil.

4m
p. 81

4.6. Strengthening National M&E Capacity

Evaluations of development programmes have historically been driven and designed by donors - primarily to satisfy their own accountability needs. However, it is increasingly recognised that both monitoring and evaluation should be a country-led process, not least because country ownership is a major factor in determining whether evaluation findings are then used in a national context.

Technical barriers to country-led evaluations centre around lack of human and financial resources, but M&E is also a highly political issue and the incentives, or lack of incentives, for evaluations to be conducted (e.g. fear of aid being withdrawn as a result of negative evaluation results) also need to be considered.

There is evidence that capacity development programmes work best when adapted to the local governance structure, professional capacity and evaluation culture.

20

> Schiavo-Campo, S., 2005, 'Building Country Capacity for Monitoring and Evaluation in the Public Sector: Selected Lessons of International Experience', World Bank Evaluation Capacity Development Working Paper no. 13, World Bank, Washington, D.C.
>
> The Evalution Capacity Development (ECD) unit of World Bank's Operations Evaluation Department is designed to help countries strengthen their monitoring and evaluation (M&E) capacity. The unit targets 'high-intensity' support to Uganda and Egypt and various other types of support to an additional 32 countries. This paper from the World Bank collates some of the main lessons learned from ECD activities and outlines the major issues which need to be addressed.

4n
p. 83

5. Participatory tools and approaches

Monitoring and evaluation (M&E) has typically been led by outside experts, measuring performance against pre-set indicators and using procedures and tools designed without the participation of key stakeholders such as the programmes' intended beneficiaries. Evaluations in particular, because they are very often conducted by external consultants, can be seen as a form of control.

There is widespread recognition that M&E should take a more inclusive, participatory approach. Participation in this sense means stakeholders are involved in deciding how the project or programme should be measured, in identifying and analysing change, and in acting on results. Nevertheless, there are few empirical studies of the effectiveness, quality and objectivity of participatory approaches. How to operationalise participation and which methods work in which contexts is the subject of ongoing debate.

5.1. The Case for Participatory M&E

Proponents of participatory monitoring and evaluation (PM&E) argue it is more ethical, cost-effective and accurate in determining the effects of interventions on people's lives than conventional approaches. Participation in decision-making processes can also motivate people to want to see those decisions implemented effectively. Another motivation for PM&E is to strengthen organisational and institutional learning.

Some argue that participatory approaches can generate both quantitative and qualitative data equally effectively.

> Chambers, R., and Mayoux, L., 2003, 'Reversing the Paradigm: Quantification and Participatory Methods', Enterprise Impact, London
>
> What role should participatory methods play in assessing the impact of development activity? A common assumption is that rigorous quantitative data can only be generated by questionnaire surveys or scientific measurement. Another is that participatory methods can only generate qualitative insights. This paper from the Enterprise Development Impact Assessment Information Service (EDIAIS) discusses experiences and innovations which show these assumptions to be false. It argues that participatory approaches can generate accurate

5a
p. 84

21

qualitative and quantitative information and should form the basis for monitoring, evaluation and impact assessment.

5.2. Challenges in Using Participatory Approaches

Whilst the ideological case for participation is widely acknowledged, PM&E is not without its critics. Crucially, labelling M&E as 'participatory' doesn't necessarily guarantee that all stakeholder groups have participated, and there are often issues around who participates and who is excluded from these processes. Subsequently, the representativeness of the findings and recommendations of participatory evaluations have been criticised.

Gregory, A., 2000, 'Problematizing Participation: A Critical Review of Approaches to Participation in Evaluation Theory', Evaluation, Vol 6, no. 2, pp. 179–199

It is widely accepted that evaluation is a social process which implies the need for a participatory approach. But what is understood by 'participation'? This critical review from Hull University Business School argues that the blanket use of the term has masked the heterogeneity evident in its realisation in practice and highlights a lack of transparency in participatory methods in evaluation.

5b
p. 86

Operationalising PM&E can also be problematic. There is a need to be sensitive to the local socio-economic and political situation, and consider under what conditions PM&E approaches can be used without increasing the vulnerabilities of already marginalised groups.

Estrella, M., 2000, 'Learning from Change: Issues and Experiences in Participatory Monitoring and Evaluation', International Development Research Centre, Ottawa

Since the 1980s concepts of participatory monitoring and evaluation (PM&E) have entered the policy-making domain of larger donor agencies and development organisations. This introductory chapter from Learning from Change: Issues and Experiences in Participatory Monitoring and Evaluation draws on twelve case studies to describe how different stakeholders have applied PM&E approaches across a range of purposes and contexts. It outlines some of the key concepts and differences between participatory and conventional approaches to M&E and highlights some emerging issues.

5c
p. 87

Vernooy, R., Qui, S., and Jianchu, X., 2006, 'The Power of Participatory Monitoring and Evaluation: Insights from South-West China', Development in Practice, Volume 16, Number 5, pp. 400-411

Participatory monitoring and evaluation (PM&E) is a partnership between researchers and other stakeholders to systematically assess research or development activities. Focusing on participatory field research for community-based natural-resource management (CBNRM), this article from the journal Development in Practice describes the capacity-building experiences of two

5d
p. 89

22

research teams in Yunnan and Guizhou provinces in south-west China. The ongoing democratisation and decentralisation processes in China aim to allow more space for local voice and decision-making power over NRM. So, who participates and what difference does participation make?

Case Study: Participatory Monitoring and Evaluation of the Zambia Social Investment Fund (ZAMSIF)

This case study provides an example of how to set up a PM&E framework.

ITAD, 2004, 'Participatory Monitoring and Evaluation: Documentation to Support Generic Framework', ITAD, London

5e
p. 91

The Participatory Monitoring and Evaluation (PM&E) consultancy for the Zambia Social Investment Fund (Zamsif) aims to establish sustainable PM&E processes at both the community and district levels. This report from ITAD and RuralNet Associates discusses the development and rationale of the PM&E framework chosen for the project. It constructs detailed diagrams to model the consultancy's work and explains current and potential uses of the framework.

5.3. Participatory Tools: Case Studies

Despite growing interest in the subject, there is no single definition or methodology of PM&E, and it encompasses a wide range of tools and approaches. Common among these approaches are values such as shared learning, democratic processes, joint decision making, co-ownership, mutual respect and empowerment. Below are some examples of participatory tools and how they have been used in practice.

5.4. Participatory Rural Appraisal

Participatory Rural Appraisal (PRA) encompasses a broad range of methods to enable local people to analyse their own realities as the basis for planning, monitoring and evaluating development activities. PRA uses group exercises to facilitate information sharing, analysis, and action among stakeholders.

Chambers, R., 2007, 'From PRA to PLA and Pluralism: Practice and Theory', IDS Working Paper, no. 286, Institute of Development Studies, Brighton

5f
p. 92

How have the theory and practice of participatory methodologies in development activities changed since the mid 1970s? What variants and applications of these methodologies have emerged? This paper from the Institute of Development Studies (IDS) traces the spread of Participatory Rural Appraisal (PRA) and Participatory Learning and Action (PLA), and identifies strengths and weaknesses in the ways they have been implemented in development contexts. Whilst inflexible applications of PRA and PLA may produce disappointing results, when executed with spontaneity and creativity, these approaches can be a source of community empowerment.

5.5. Key Informant Interviews

Key informant interviews are a rapid assessment methodology which can be used as an intermediate indicator of outcomes as an alternative or supplement to full impact assessments.

Price, N., and Pokharel, D., 2005, 'Using Key Informant Monitoring in Safe Motherhood Programming in Nepal', Development in Practice, Volume 15, no. 2, pp 151-164

The Nepal Safer Motherhood Project (NSMP) works to improve maternal health and contribute to programme development at district and national level. This article from the journal Development in Practice discusses the project's use of Key Informant Monitoring (KIM). KIM is an adapted version of the peer ethnographic research method. Data is collected by community-based Key Informant Researchers (KIRs) and used for monitoring and planning. KIRs have proved useful sources of information and acted as change agents by spreading safer motherhood messages.

5g
p. 94

5.6. Most Significant Change Technique

The Most Significant Change (MSC) technique involves the collection of change stories emanating from the field level, and the systematic selection of the most significant of these. These selected stories are then discussed and critically reflected on to help determine the impact of the development programme or activity.

Davies, R., 2005, 'The 'Most Significant Change' (MSC) Technique: A Guide to Its Use', MandE, London

This paper from MandE News, an online monitoring and evaluation news service, outlines an innovative qualitative monitoring technique known as the 'most significant change' (MSC) approach. The MSC technique is a participatory method of collecting and analysing stories from the field which focuses on monitoring intermediate outcomes and impact. It provides a simple means of making sense of a large amount of complex information and is best suited to large-scale, open-ended projects which would otherwise be difficult to monitor using traditional methods.

5h
p. 95

Case Study: Cabungo in Malawi

There is evidence that the most significant change technique can enhance organisational learning and performance.

Wrigley, R., 2006, 'Learning from Capacity Building Practice: Adapting the 'Most Significant Change' (MSC) Approach to Evaluate Capacity Building', INTRAC Praxis Paper no. 12, International NGO Training and Research Centre, Oxford

There is growing recognition of the need to take a multi-stakeholder approach to evaluation, which promotes local ownership and builds capacity for reflection, learning, improved performance and self-determination. This paper from the International NGO Training and Research Centre (INTRAC) reflects on the use of the 'Most Significant Change' (MSC) methodology to evaluate the capacity

5i
p. 97

24

building services of CABUNGO, a local capacity building support provider in Malawi.

5.7. Outcome Mapping

Outcome Mapping is an alternative to theory-based approaches to evaluation that rely on a cause–effect framework; rather, it recognizes that multiple, nonlinear events lead to change. It focuses on people and changes of behaviour and how far development interventions have built the capacity of the local community. Outcome mapping assumes only that a contribution has been made, and never attempts attribution.

Earl S., Carden, F., and Smutylo., T., 2001, 'Outcome Mapping: Building Learning and Reflection into Development Programs', International Development Research Centre, Ottawa

5j
p. 99

Development organisations are increasingly under pressure to demonstrate that their programmes result in positive changes. This paper from the International Development Research Centre (IDRC) argues that impacts are often the product of events for which no single agency can claim full credit. Outcome mapping moves away from impact assessments to focus on changes in the behaviour of the people it works with directly.

6. Gender and conflict sensitivity

6.1. Gender Analysis

Development interventions can have differential impacts on men and women: Men and women have different needs and constraints, different opportunities to participate in programme design and implementation, and benefit differently from outcomes and impacts. A gender analysis framework should therefore be a component of all evaluation designs. It is also important to have a gender balanced evaluation team.

World Bank, 2002, 'Chapter 10: Integrating Gender into the PRS Monitoring and Evaluation, ' in PRSP Sourcebook, World Bank, Washington, D.C.

6a
p. 100

There is growing evidence that gender-sensitive development strategies contribute significantly to economic growth and equity objectives by ensuring that all groups of the poor share in programme benefits. Yet differences between men's and women's needs are often not fully recognised in poverty analysis and participatory planning, and are frequently ignored in the selection and design of Poverty Reduction Strategies (PRSs). A full understanding of the gender dimensions of poverty can significantly change the definition of priority policy and programme interventions supported by the PRS. This chapter from the World Bank PRSP Sourcebook provides practical guidance on identifying and implementing country-level policies and programmes that will benefit both men and women, and maximise potential benefits for poor families.

Donor agencies have had limited success in introducing gender sensitive approaches to M&E. This is particularly evident in the use of household surveys which either only

interview the household head (usually male) or which interview women in contexts where they are not able to speak freely.

Evaluations should address critical gender issues such as time poverty, participation in household decision-making and women's multiple roles (e.g. production, social reproduction and community management).

Bambrilla, P., 2001, 'Gender and Monitoring: A Review of Practical Experiences', report prepared for the Swiss Agency for Development Cooperation, BRIDGE Report 63, BRIDGE, Institute of Development Studies, Brighton

6b
p. 102

How can monitoring and evaluation (M&E) processes be made gender-sensitive? What measures have organisations taken to assess their effectiveness in mainstreaming gender? This report by BRIDGE for the Swiss Agency for Development and Cooperation (SDC) provides a tool for integrating a gender approach into M&E mechanisms. It draws on a number of case studies and different experiences of organisations implementing M&E mechanisms and provides a number of recommendations for implementing gender-sensitive mechanisms.

6.2. M&E in Conflict and Fragile States

M&E faces unique challenges in situations of extreme uncertainty, particularly in terms of information gathering. These include:

- Lack of security (both for researchers and the people who talk to them)
- Distrust of outsiders and reluctance to talk, or fear of reprisals for talking
- Shame in acknowledging sexual violence
- Rapidly fluctuating populations (due to migration) and no adequate documentation on the population
- Government's unwillingness to recognise the level of violence or the failure of programmes which might in turn affect aid flows.

Elkins, C., 2006, 'Monitoring and Evaluation (M&E) for Development in Peace-Precarious Situations', RTI International, London

6c
p. 103

How can monitoring and evaluation (M&E) information systems improve programme impact and assist peaceful development in situations prone to violent conflict? This paper from RTI International outlines M&E's status as a unique discipline and describes M&E strategies and tactics implemented in real-world, peace-precarious situations. Even under the stresses of violence and conflict, M&E approaches can help build knowledge of how to push forward peace and development.

In fragile environments, choosing appropriate tools to monitor and evaluate programmes is a key concern.

> Social Impact, 2006, 'Monitoring, Evaluating and Learning for Fragile States and Peacebuilding Programs, Practical Tools for Improving Program Perfomance and Results', guide prepared for USAID, Social Impact, Arlington, VA
>
> How can organisations implement fragile states peacebuilding (FSP) programmes with realistic development outcomes that can rapidly adapt to changing circumstances? This guide from Social Impact aims to increase the effectiveness of FSP programmes through more systematic approaches to Monitoring, Evaluation and Learning (ME&L). Stronger ME&L enables agencies and communities to understand what is effective, how to consolidate best practice and how to increase accountability to stakeholders.

6d
p. 105

Case Study: M&E in Nepal

In DFID Nepal, monitoring and evaluation involves closely monitoring local context, and giving consideration to the impact of programming on the conflict and on social inclusion.

> Department for International Development, 2005, 'Nepal Country Assistance Plan: Monitoring in a Fragile State', DFID, London
>
> The DFID Nepal Country Assistance Plan (CAP), published in February 2004, aims to reduce poverty and social exclusion and help establish the basis for lasting peace. To this end, CAP commits DFID Nepal to developing improved, locally accountable and transparent systems for monitoring progress. This DFID paper describes the main elements of the CAP Monitoring Framework (CAPMF), setting out how DFID Nepal (DFIDN) proposes to monitor the operating environment for development activities and measure progress towards outputs.

6e
p. 106

7. Appendix: Full summaries

1a. Monitoring and Evaluation: Some Tools, Methods and Approaches

Author: World Bank Operations Evaluation Department
Date: 2002
Size: 23 pages

Monitoring and Evaluation (M&E) is an area of growing importance for the development community. It allows those involved in development activities to learn from experience, to achieve better results and to be more accountable. This report from the World Bank Operations Evaluation Department provides an overview of some of the M&E tools, methods and approaches on offer to development practitioners.

There is increased interest in M&E among the development community due to a stronger focus on the results produced by interventions. M&E processes allow those involved to assess the impact of a particular activity, to determine how it could be done better and to show what action is being taken by different stakeholders. This should translate into a more effective and transparent way of working. However, there is some confusion about what M&E entails – a problem which the report aims to help solve. It details nine different M&E tools and approaches, including data collection methods, analytical frameworks and types of evaluation and review. For each, it lists their purpose and use; advantages and disadvantages; the required costs, skills and time; and key references. The choice of which to use in a given context will depend on considerations such as the purpose for which M&E is intended, the main stakeholders, how quickly the information is needed and the cost.

It is emphasised that the list is not comprehensive. Some of the tools and approaches are complementary or substitutes; some are broad in scope, others narrower. The nine M&E methods are:

- Performance indicators. These measure inputs, processes, outputs, outcomes and impacts of development interventions. They are used for setting targets and measuring progress towards them.
- The logical framework (LogFrame) approach. This identifies objectives and expected causal links and risks along the results chain. It is a vehicle for engaging partners and can help improve programme design.
- Theory-based evaluation. Similar to the LogFrame approach, this provides a deeper understanding of the workings of a complex intervention. It helps planning and management by identifying critical success factors.
- Formal surveys. These are used to collect standardised information from a sample of people or households. They are useful for understanding actual conditions and changes over time.

28

- Rapid appraisal methods. These are quick, cheap ways of providing decision-makers with views and feedback from beneficiaries and stakeholders. They include interviewing, focus groups and field observation.

- Participatory methods. These allow stakeholders to be actively involved in decision-making. They generate a sense of ownership of M&E results and recommendations, and build local capacity.

- Public expenditure tracking surveys. These trace the flow of public funds and assess whether resources reach the intended recipients. They can help diagnose service-delivery problems and improve accountability.

- Cost-benefit and cost-effectiveness analysis. These tools assess whether the cost of an activity is justified by its impact. Cost-benefit measures inputs and outputs in monetary terms, whereas cost-effectiveness looks at outputs in non-monetary terms.

- Impact evaluation. This is the systematic identification of the effects of an intervention on households, institutions and the environment, using some of the above methods. It can be used to gauge the effectiveness of activities in reaching the poor.

Source: World Bank Operations Evaluation Department, 2002, Monitoring and Evaluation: Some Tools, Methods and Approaches, The World Bank, Washington, D.C
Author: Independent Evaluations Group, World Bank , eline@worldbank.org
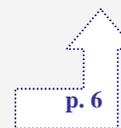
Full text available online:
http://lnweb18.worldbank.org/oed/oeddoclib.nsf/24cc3bb1f94ae11c85256808006a0046/a5efbb5d776b67d285256b1e0079c9a3/$FILE/MandE_tools_methods_approaches.pdf

---

1b. Ten Steps to a Results-based Monitoring and Evaluation System

Author: J Kusek and R Rist
Date: 2004
Size: 268 pages

Governments and organisations face increasing internal and external pressures to demonstrate accountability, transparency and results. Results-based monitoring and evaluation (M&E) systems are powerful public management tools to achieve these objectives. This handbook from the World Bank presents a ten-step model that provides extensive detail on building, maintaining and sustaining a results-based M&E system.

Results-based M&E systems can help build and foster political and financial support for policies, programmes and projects and can help governments build a solid knowledge base. They can also produce major changes in the way governments and organisations operate, leading to improved performance, accountability, transparency, learning, and knowledge. Results-based M&E systems should be considered a work in progress. Continuous attention, resources, and political commitment are needed to ensure their

viability and sustainability. Building the cultural shift necessary to move an organisation toward a results orientation takes time, commitment and political will.

The ten steps to building, maintaining and sustaining a results-based M&E system are outlined below:

- A readiness assessment should be conducted to determine whether prerequisites for a results-based M&E system are in place. It should review incentives and capacity for an M&E system and roles, responsibilities and structures for assessing government performance.

- Outcomes to monitor and evaluate should be agreed through a participatory process identifying stakeholders' concerns and formulating them as outcome statements. Outcomes should be disaggregated and a plan developed to assess how they will be achieved.

- Key performance indicators to monitor outcomes should be selected through a participatory process considering stakeholder interests and specific needs. Indicators should be clear, relevant, economical, adequate and monitorable.

- Baseline data on indicators should be established as a guide by which to monitor future performance. Important issues when setting baselines and gathering data on indicators include the sources, collection, analysis, reporting and use of data.

- Performance targets should be selected to identify expected and desired project, programme or policy results. Factors to consider include baselines, available resources, time frames and political concerns. A participatory process with stakeholders and partners is key.

- Monitoring for results includes both implementation and results monitoring, as well as forming partnerships to attain shared outcomes. Monitoring systems need ownership, management, maintenance and credibility. Data collection needs reliability, validity and timeliness.

- Evaluation provides information on strategy, operations and learning. Different types of evaluation answer different questions. Features of quality evaluations include impartiality, usefulness, technical adequacy, stakeholder involvement, value for money and feedback.

- Reports on the findings of M&E systems can be used to gain support and explore and investigate. Reports should consider the requirements of the target audience and present data clearly.

- Findings of results-based M&E systems can also be used to improve performance and demonstrate accountability and transparency. Benefits of using findings include continuous feedback and organisational and institutional knowledge and learning.

- Good results-based M&E systems must be used in order to be sustainable. Critical components of sustaining M&E systems include demand, clear roles and responsibilities, trustworthy and credible information, accountability, capacity and incentives.

Source: Kusek, J., and Rist, R., 2004, 'Ten Steps to a Results-based Monitoring and Evaluation System', World Bank, Washington, D.C.
Author: Jody Zall Kusek and Ray C. Rist, World Bank, www.worldbank.org
⬤ Full text available online: http://www.oecd.org/dataoecd/23/27/35281194.pdf

2a. Putting the New Aid Paradigm to Work: Challenges for Monitoring and Evaluation

Author: N Holvoet and R Renard
Date: 2006
Size: 41 pages

What are the challenges in monitoring and evaluating Poverty Reduction Strategy Paper (PRSP) outcomes? Have issues been neglected in these key areas? This study by the Institute of Development Policy and Management, University of Antwerp, reviews and assesses M&E systems for 11 Sub-Saharan African countries. These systems are embryonic in the countries reviewed, and are a weak link in most PRSPs. The emphasis on participation and comprehensiveness places unrealistic demands on national capacity, exacerbated by the scant attention given to M&E by donors.

The PRSP approach is the new paradigm for providing effective aid. It replaces donor-imposed instruments with country-produced strategies. These address social, environmental and governance as well as macroeconomic issues, and entail broad based civil society involvement. Sound M&E systems are crucial to the success of the PRSP approach, but have been largely overlooked by national and donor policy makers. No clear cut distinction is made between monitoring and evaluation, or the linkages between them. This results in loss of the balance between accountability and feedback. Evaluation is treated as an afterthought, with harmful effects on data acquisition and feedback to national policies. The PRSP rationale requires M&E systems that are multi-stakeholder, multi-purpose, multi-dimensional, multi-method, multi-layer, and multi-criteria. These requirements are daunting for the embryonic M&E systems in PRSP countries. Furthermore:

- Recipient countries have weak public sectors and limited human resource capacity which cannot cope with the complex tasks of M&E
- M&E design is often poor. There is too much attention on inputs and final poverty outcomes leading to a 'missing middle' in identifying intermediary outcomes.
- Problems with data quality, timeliness, and unreliable reporting mechanisms are not taken into account in identifying indicators
- Participation is limited to preparation stages of the PRSP and largely absent in implementation. There is little shared control in decision making.
- Information from participatory assessments and conventional surveys is not integrated.
- There is a dearth of analysis once indicators have been collected.

31

There is a pressing need for more attention to be given to M&E in order to cope with the many simultaneous challenges it represents. The national planning, budgeting, and M&E capacity required by PRSPs force institutional and structural issues onto the reform agenda. Progress in achieving these is difficult and long-term, but PRSPs are medium term strategies. Donors need to look at the timeframes for replacing their M&E systems with national ones. This is essential if the emphasis is on outcomes rather than aid spending ratios. Given the timeframe constraints, attention must be paid to getting satisfactory results from systems under construction. It is also necessary to:

- engage Parliament in M&E, especially in relation to accountability. Donors have neglected this important dimension.

- ensure that M&E meets the information needs of donors as well as governments and national stakeholders. Failure to do this will result either in parallel systems or increased burdens on national ones.

- give line ministries a central role in monitoring, with shared responsibility for evaluation with central agencies.

- monitor and evaluate the overall policy chain. This will address data unreliability and integration issues. It entails having a clear 'programme theory' as the conceptual base for M&E.

- integrate Public Expenditure Management with M&E to clarify how indicators flow from the PRSP.

Source: Holvoet, N., and Renard, R., 2006, 'Putting the New Aid Paradigm to Work: Challenges for Monitoring and Evaluation', Institute of Development Policy and Management, Antwerp
Author: Nathalie Holvoet and Robrecht Renard, Institute of Development Policy and Management, http://www.ua.ac.be/dev
Full text available online:
http://www.ua.ac.be/download.aspx?c=.IOBE&n=37390&ct=38167&e=85310

2b. Monitoring and Evaluation: Chapter 3 in PRSP Sourcebook

Author: G Prennushi and G Rubio
Date: 2002
Size: 26 pages

How do we know if a poverty reduction strategy is effective? First, a poverty monitoring system is needed to track key indicators over time and space and to determine if they change as a result of the strategy. Second, rigorous evaluations should be done selectively to assess the impact on poverty of interventions that are key components of the strategy. This chapter from the World Bank's *Poverty Reduction Strategy Paper Sourcebook* examines the features of poverty monitoring systems and explores some of the key issues which arise during implementation.

The Millenium Development Goals (MDGs) provide a prime example of the types of goals, indicators and targets that can be used to monitor progress. In this context, goals are the objectives a country or society wants to achieve. Indicators are the variables used to measure progress. Targets are the quantified levels of the indicators that a country wants to achieve at a given time.

Some of the key challenges and features of a poverty monitoring system are outlined below:

- Many countries already have poverty monitoring systems in place so the task is to assess their adequacy and strengthen them as necessary.

- Rigorous evaluations should be done selectively to assess the impact on poverty of key interventions. Other types of evaluation, such as assessing the process of formulating a poverty reduction strategy, can also be useful.

- Another challenge is to evaluate the impact of poverty reduction strategies in general as opposed to specific components of a strategy.

- Much monitoring and evaluation takes place without adequate development of in-country capacity and without strong links to the key decision-makers. Precious opportunities to learn what works are lost, sometimes along with funds.

- Results that are not widely disseminated, through mechanisms tailored to different groups in civil society, will not be used, and the resources spent in getting such results will be wasted.

- Nongovernmental actors have a key role to play in the design of monitoring and evaluation systems, and in their implementation and in using results.

Policy officials setting up a poverty monitoring system should be aware of a number of issues:

- The tracking of public expenditures and outputs and quick monitoring of household well-being need special attention. Participatory data collection methods and qualitative information should not be overlooked.

- Impact evaluations can be demanding in terms of analytical capacity and resources. It is important that they are conducted only when the characteristics of the intervention warrant an evaluation.

- The dissemination strategy should accommodate the diverse information needs of different groups. Information should be tailored to the relevant audience: press releases for the media and workshops and seminars for the general public and civil organisations.

- A well established process to feed monitoring and evaluation results back to policymakers is crucial if results are to be used in formulating policy. Since key decisions are made at the time of budget formulation, key results should be available then.

- Broad consultations are required during the design of the monitoring and evaluation system to build consensus on what to monitor and evaluate. They

generate a sense of ownership among different groups in society, increasing acceptance and use of findings.

Source: World Bank, 2002, 'Chapter 3: Monitoring and Evaluation', in PRSP Sourcebook, World Bank, Washington, D.C.
Author: World Bank , www.worldbank.org
Full text available online: http://povlibrary.worldbank.org/files/4480_chap3.pdf

---

2c. Research on the Current State of PRS Monitoring Systems

Author: H Lucas et al
Date: 2004
Size: 70 pages

p. 8

What are the key factors that determine the success or failure of Poverty Reduction Strategies (PRS) monitoring efforts? Who should be involved? How should the information gathered be put to use? This study by the Institute of Development Studies reviews the main issues arising from the implementation of PRS monitoring systems. It draws on the experience of different countries and suggests possible monitoring approaches to suit specific contexts.

Success or failure of PRS monitoring depends on the personality, status and capabilities of a few key players. Involving key stakeholders is essential to ensure their contribution to the monitoring of the PRS. Although government views are reflected in PRS Papers, the priorities of ministries and other agencies might not be. Creating a technical secretariat with adequate capacity to carry out the basic and more detailed monitoring activities can reduce the burden on other agencies and help get their support. Labelling the monitoring process as a 'high-status' initiative can also ensure that key individuals and ministries are willing to be seen backing this effort.

Other relevant findings of the study are outlined below:

- High-level committees can be created to bring key players on board. Such committees can play a coordinating and oversight role and leave the more testing activities to the technical secretariat.

- Civil Society Organisations (CSOs) need to be adequately involved in the process. If this goes beyond merely informing them of government plans, their involvement can add transparency and accountability to the monitoring effort.

- Parliaments have played a minimal role in the monitoring of PRS. Yet, their involvement is a key component of a truly democratic framework.

- Decentralisation processes need to be accompanied by local, as well as national, monitoring systems. Yet, whereas lack of capacity is a key constraint at national level, the problem is even more acute at local level.

- 'Missing middle indicators', which refer to how well intentioned policies will deliver promised outcomes/impacts, continue to be a concern. Given scarce resources, focusing on budget allocations and expenditure, as well as making better use of Routine Data Systems might be a possible solution.

Monitoring data is still primarily produced to meet donor requirements. Governments do not seem interested in developing data for their own purposes. In this context, donors need to:

- ensure that a broad range of stakeholders are involved in the monitoring efforts. This can contribute to country ownership of the process.

- avoid linking performance monitoring and donor funding as guiding principles for PRS implementation. Incentives to monitor can decrease if the identification of problems leads to funding cuts or termination.

- introduce specific incentives to ensure that monitoring data informs policy and budgetary policies. Decisions continue to be based on political and personal interests.

- build capacity within the PRS monitoring agencies to communicate and share monitoring data in a way that matches the needs and interests of different stakeholders.

- support the development of comparative and ranking indicators at various levels. This can facilitate dissemination and policy influence. It can also help persuade governments and CSOs of the value of the data produced.

Source: Lucas, H., Evans, D. and Pasteur, K., 2004, 'Research on the Current State of PRS Monitoring Systems', IDS Discussion Paper no. 382, Institute of Development Studies, Brighton
Author: Henry Lucas, David Evans and Katherine Pasteur, Institute of Development Studies , http://www.ids.ac.uk/ids
Full text available online: http://www.ids.ac.uk/ids/bookshop/dp/dp382.pdf

---

2d. Good Practice in the Development of PRSP Indicators and Monitoring Systems

Author: D Booth and H Lucas
Date: 2002
Size: 73 pages

p. 8

How can poverty reduction strategies best be monitored? And how can monitoring lead to greater success in reducing poverty? This study by the Overseas Development Institute (ODI) draws conclusions about best practice from a review of Poverty Reduction Strategy Papers (PRSPs) and suggests ways in which monitoring can be improved. PRSP monitoring calls for fresh thinking. It needs to be geared to what is new and challenging about the PRSP initiative – particularly the effort to engage a wider range of stakeholders in policy dialogue about poverty reduction at the national level. It also needs an

35

understanding of the relevant policy processes and the possible uses of information in enforcing new kinds of accountability and learning about poverty reduction.

Furthermore, the greater results-orientation that is a feature of the PRSP approach should not be taken to imply an exclusive interest in monitoring final results or impacts. PRSPs are leading to a major upsurge in final poverty-outcome measurement. However, there is much less evidence of renewed interest in measuring the intermediate processes and achievements that will be necessary to produce the desired final outcomes. This is a serious deficiency, as rapid feedback on this level of change is what matters most for accountability and learning. The poor quality of the administrative reporting systems on which much of the relevant data depend is being ignored.

PRSPs also pay little attention to the possibility of using alternative methods to compensate for the unreliability of routine information systems. Furthermore, it is unclear how stakeholders will be incorporated into PRSP monitoring arrangements, and generally how information will be used to improve policy and implementation. Some of the key findings are outlined below:

- A multidimensional approach to final poverty outcomes is increasingly accepted but still poses significant challenges
- Knowing which intermediate variables to monitor is not easy. The selection needs to involve strategic thinking
- Tracking financial and non-financial inputs can lead to policy improvements that are important for poverty reduction
- Despite its aura of technical superiority, survey-based analysis of poverty trends can be very inaccurate
- Improvements in routine information systems are possible, but they call for both realism and a very imaginative approach
- Service-delivery surveys, problem-oriented commissioned studies and participatory impact monitoring (PIM) have been proven to be useful complements to administrative data.

In order for PRSP monitoring to lead to an improvement in poverty reduction, the following should be taken into account:

- It is useful to distinguish between long-term institutional solutions to the lack of demand for poverty-related information, and worthwhile interim measures
- Successful arrangements are likely to be those that are well supported politically and also permit swift executive action when necessary.

Source: Booth, D. and Lucas, H. 2002, 'Good Practice in the Development of PRSP Indicators', Overseas Development Institute (ODI) Working Paper 172, ODI, London
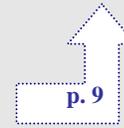**Author:** David Booth and Henry Lucas, ODI, http://www.odi.org.uk/
Full text available online: http://www.odi.org.uk/publications/working_papers/wp172.pdf

2e. Poverty Monitoring Guidance Note 1: Selecting Indicators

Author: World Bank
Date: 2004
Size: 5 pages

What is the best method for selecting appropriate indicators to monitor the implementation of poverty reduction strategies? This paper from the World Bank provides a summary of good practice for selecting such indicators. In general, it is preferable to select a few good quality indicators which are easily measurable within the current capacity and which covering the right questions, at the right level of disaggregation. This is ultimately a political process which needs to take into account existing constraints.

Once a goal has been set for a poverty reduction strategy, it is critical to monitor indicators at various stages of programme implementation in order to understand where changes or additional efforts may be needed. In the light of political and resource constraints, it is important to ensure consistency by identifying and revising data priorities at early stages of planning.

Indicators can be grouped into two categories: intermediate (covering inputs into and outputs from policies and programmes) and final (highlighting outcomes and impacts on households and individuals):

- Intermediate indicators measure factors that contribute to the process of achieving an outcome or impact. These include input indicators, which measure the various financial and physical resources dedicated to a goal, and output indicators, which measure the goods and services that are produced by the inputs.

- Final indicators measure the effect of an intervention on individuals' well-being. These include outcome indicators, which assess the use of, and satisfaction with, public services, and impact indicators, which measure key dimensions of well-being such as literacy or health.

- Monitoring final indicators helps judge progress towards set goals. Monitoring intermediate indicators allows ongoing observation of the changing dynamics of policy and programming implementation.

- Aggregate country-level indicators are useful for giving an overall picture. However, they tend to mask significant policy-relevant differences between groups. Indicators can be disaggregated along various dimensions, including location, gender, income level, and social group, in order to give a more detailed view.

- The selection and disaggregation of indicators usually has political consequences insofar as they reflect priorities and induce accountability. Furthermore, the choice of indicators depends on the types of data available in a country.

Good quality indicators share a number of features which can be used as a checklist when selecting the most relevant factors to measure. Such indicators should:

- be a direct, unambiguous measure of progress
- vary across groups, areas, and over time
- have a direct link with interventions
- be relevant for policy making
- be consistent with the decision-making cycle
- not be easily influenced by unrelated developments
- be easy and not too costly to measure
- be easy to understand and reliable
- be consistent with data availability.

Source: World Bank, 2004, 'Poverty Monitoring Guidance Note 1: Selecting Indicators', World Bank, Washington, D.C.
Author: World Bank , www.worldbank.org
Full text available online:
http://siteresources.worldbank.org/INTPAME/Resources/Selective-Evaluations/NoteIndicators_eng_Apr04_doc.pdf

---

2f. Measuring Democratic Governance: A Framework for Selecting Pro-poor and Gender Sensitive Indicators

Author: United Nations Development Programme
Date: 2006
Size: 60 pages

p. 9

How should we measure democratic governance? Most indicators are developed by external stakeholders to compare nation states and are not designed to help countries undertake governance reforms. This guide from the United Nations Development Programme (UNDP) presents a framework for generating pro-poor gender sensitive indicators to help policy makers monitor and evaluate democratic governance at the country level. It argues that indicator selection is itself a governance process.

Measuring democracy is a complex task. Indicators help show how much progress is being made towards goals set out in national development plans. Even when indicators are developed by national stakeholders, they often fail to focus on poorer groups and the different experiences of men and women. The Democracy Assessment Framework from The International Institute for Democracy and Electoral Assistance (IDEA) is a useful source of basic principles. From it, key democratic values can be drawn: participation, representation, accountability, transparency, responsiveness, efficiency and equity.

There are four senses in which indicators may be considered pro-poor or gender specific:

- Disaggregated by poverty status or sex. For example, the ratio of male to female voters

38

- Specific to the poor or gender specific. For example, the proportion of cases brought to trial from non-poor households

- Implicitly pro-poor or gendered. For example, hours per day that polling booths are open (opportunity for labourers and shift workers to vote)

- Chosen by the poor or chosen separately by men and women. For example, the percentage of women who say they are given adequate information on their rights.

Three tools can be used to assess the need for pro-poor and gender sensitive indicators:

- A set of key questions themed around areas of governance

- A process flow chart showing steps within legal, administrative and political processes

- An integrated indicator matrix providing an overall picture of where gender sensitive and pro-poor indicators are needed.

The following points are important in relation to the implementation of this framework:

- The framework can be applied to seven areas of governance: parliamentary development, electoral systems and processes, human rights, justice, access to information and the media, decentralisation and local governance, and public administration reform and anti-corruption.

- Countries where no poverty monitoring system is in place can follow a timetable of activities leading up to indicator selection. This begins with an announcement of the intention to develop pro-poor gender sensitive ways of measuring governance.

- Substantial data is needed to support pro-poor and gender sensitive indicators. This may be from single or multiple sources including surveys, administrative data and qualitative methods. Second-generation indicators promise to improve the quality of statistics.

- These tools can only be used to promote pro-poor and gender sensitive governance if they are supported by a range of national stakeholders. Inclusive and participatory debate should inform indicator selection and processes of data collection and evaluation.

Source: United Nations Development Programme, 2006, 'Measuring Democratic Governance: A framework for selecting pro-poor and gender sensitive indicators', UNDP, Oslo
Author: Christopher Scott and Alexandra Wilde, UNDP: www.undp.org
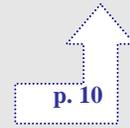 Full text available online:
http://www.undp.org/oslocentre/docs06/Framework%20paper%20-%20entire%20paper.pdf

2g. Beyond the Numbers: Understanding the Institutions for Monitoring Poverty Reduction Strategies

Author: T Bedi et al
Date: 2006
Size: 260 pages

Monitoring systems are central to the effective design and implementation of a Poverty Reduction Strategy (PRS). Many existing PRS monitoring systems lack coordination and a coherent institutional framework linking monitoring and decision making. This World Bank report draws on 12 country studies to conclude that PRS monitoring systems should build on existing elements to begin a process of gradual change. Clearly defined relationships, incentives and activities and identification of entry points in decision-making processes facilitate the supply of monitoring information and its integration into improving PRS policies.

In most countries, both the supply side (organising monitoring and reporting of indicators across fragmented administrations) and the demand side (ensuring that monitoring information is used in decision-making) are posing major practical challenges. Inconsistent monitoring frameworks cause duplication, excessive administrative burdens, lack of data compatibility and poor information flows. The main challenges are rationalising existing monitoring mechanisms and coordinating numerous separate actors.

A PRS monitoring system incorporates three main functions: tracking progress in poverty reduction against indicators, implementation of the monitoring and evaluation of the PRS and expenditure tracking:

- Common obstacles to effective PRS monitoring systems include: lack of operational detail, costing and prioritisation in PRSs; data coordination difficulties between agencies; capacity constraints in data collection; underdeveloped budget and expenditure management and barely institutionalised links between monitoring and policy-making.

- A design process that includes participatory methods and stakeholder analysis is critical to ensure buy-in from stakeholders and establish joint objectives and solutions.

- Process design should build on existing systems and activities to prevent duplication and competition. Flexibility in institutional design allows the emerging system to evolve and adapt to change.

- Strong leadership and effective advocacy from the centre of government gives PRS monitoring systems authority and facilitates linkages with policy and budget processes.

- Coordination mechanisms that minimise the burden on participants and develop incentives for participation, along with strategies for building monitoring and analytical capacity throughout government, are important.

- Aligning donor reporting requirements to the PRS monitoring system would assist coordination between different sector agencies.

Evidence-based policy making and institutional learning are key objectives of the PRS initiative, but are difficult to institutionalise. The best strategy for strengthening demand for monitoring information is to tailor system outputs to key points in the policy-making process:

- Analysis and evaluation need to be institutionalised in PRS monitoring systems. Creating central analytical units and harnessing the analytical capacity of non-governmental actors are useful techniques.

- Information and analysis must be tailored to the needs of decision-makers and users for dissemination across central and local governments, service providers, parliaments, the media, the public and donors.

- Linking monitoring to the budget and planning process requires public sector agencies to justify bids for funding according to PRS objectives and program performance.

- Involving parliaments in monitoring systems aids their ability to carry out executive oversight and represent their constituencies.

- Although the participation of civil society is largely informal and varies considerably, it can include performing monitoring activities, participating in committees and working groups, providing analysis and advice and disseminating outputs to the public.

Source: Bedi, T. et al., 2006, 'Beyond the Numbers: Understanding the Institutions for Monitoring Poverty Reduction Strategies', World Bank, Washington DC
Author: Tara Bedi, Aline Coudouel, Marcus Cox, Markus Goldstein and Nigel Thornton
The World Bank, http://www.worldbank.org
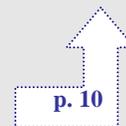Full text available online:
http://siteresources.worldbank.org/INTPAME/Resources/383220-1153403450741/0821364847.pdf

2h. Poverty Monitoring Systems: An Analysis of Institutional Arrangements in Uganda

Author: D Booth and X Nsabagasani
Date: 2005
Size: 51 pages

Monitoring activities can play an important role in promoting country ownership of poverty-reduction policies where they are closely related to a politically supported and maturing budget process. Focusing on Uganda, this paper from the Overseas Development Institute (ODI) is one of a series of country studies intended to inform the design and implementation of poverty monitoring systems (PMS). While Uganda had an early start in creating a political environment that enabled results to influence policy,

41

problems with incentives and partnership continue. There will inevitably be setbacks in developing a PMS - a process that requires co-ordination, policy-relevance and strong donor-country partnerships.

Uganda made its first attempt to define a strategy to monitor the progress of its Poverty Eradication Action Plan (PEAP) in 2001. While the Poverty Monitoring Evaluation Strategy (PMES) facilitated better reporting, it did not impose real coordination. The National Integrated Monitoring and Evaluation Strategy (NIMES), a more inclusive monitoring strategy, was formulated in 2003-2004. An early misconception that NIMES would simply be an additional monitoring system has been mostly, but not entirely, cleared up.

The successes, limitations and challenges of Uganda's PMS relate to country ownership, results orientation and aid-alignment:

- NIMES appears to be effective in strengthening country ownership of poverty monitoring. However, work still needs to be done to dispel misunderstandings about its objectives and approach.

- Systems suffer from duplication and haphazardness. The incentive to generate and use quality information is weak compared to incentives to defend and extend existing activities. With over 500 donor-funded projects in Uganda, territoriality is an issue.

- While there has been a good supply of information on many topics, better co-ordination of data collection and analysis is a key priority under NIMES.

- The inclusion of a Policy Matrix, setting out agreed policy actions, as well as a Results and Monitoring Matrix in the 2004 PEAP is a crucial innovation which offers a framework for more systematic monitoring.

- PEAP matrices also provide the opportunity for better aid alignment by reducing the gap between donor and government review processes.

- The inclusion of a Research and Evaluation Working Group in the NIMES plan is encouraging since these aspects have been previously under-emphasised in PMS.

There are a number of general lessons which emerge, as well as specific implications for Uganda:

- The coordination of monitoring should not simply be treated as a technical challenge, subject to administrative solutions. Improving coordination and promoting results-orientated policy requires attention to the incentive structure of the civil service.

- Monitoring systems also need to get better at addressing what policy makers need to know. If the PMS does not provide policy-relevant information on an appropriate timescale, it will not be relied upon by policy makers.

- A major challenge in the partnerships involved in implementing poverty reduction strategies is the gap between country PRS monitoring and donor monitoring. A convergence of approach would support domestic policy-making.

- Tensions between Uganda's Ministry of Finance, Planning and Economic Development and its new National Planning Authority are significant. These go beyond monitoring and need to be resolved at the political level.
- While a distinct research and evaluation component for NIMES is a good proposal, attention should be paid to the successes and limitations of Tanzania's experience.

Source: Booth, D., and Nsabagasani, X., 2005, 'Poverty Monitoring Systems: An Analysis of Institutional Arrangements in Uganda', ODI Working Paper No. 246, Overseas Development Institute, London
Author: David Booth and Xavier Nsabagasani, Overseas Development Institute (ODI), http://www.odi.org.uk/
🔲 Full text available online:
http://www.odi.org.uk/Publications/working_papers/wp246.pdf

2i. The Role of Donors in PRS Monitoring Systems: Advantages and Disadvantages of the Different Forms of Involvement

Author: H Lucas and M Zabel
Date: 2005
Size: 39 pages

How can donors improve the way in which they engage with the monitoring of Poverty Reduction Strategy (PRS) implementation? This draft report prepared for the World Bank by the HLSP Institute considers the reported activities and experiences of donors in PRS monitoring. It assesses the current position and provides some guidance as to possible ways forward.

The demand for PRS monitoring information is dominated by the needs of donors and findings are rarely tailored to meet the specific requirements of national stakeholders. Moreover, this demand places considerable burdens on overstretched information systems. The challenge of PRS monitoring involves the need to (a) provide up to date, reliable information; (b) encompass both PRS implementation processes and poverty outputs/outcomes; and (c) persuade donors of the quality of information. Whilst the desirability of country ownership and harmonisation and alignment is well established, there remains a substantial gap between theory and practice.

The role of donors in PRS monitoring is considered in relation to six key areas:

- Donors need to support the design and implementation of PRS monitoring systems that can track processes and provide timely, diagnostic information.
- The effectiveness of a move from project aid to budget support requires the commitment of the recipient government and a common set of objectives.

- Accountability is still predominantly viewed in terms of meeting the requirements of donors. Moreover, many key stakeholders lack clarity as to the nature of those requirements. There is also widespread uncertainty as to whether the PRS should be considered as *the* national framework for development.

- A donor-driven monitoring process may be unavoidable. The inequality of the negotiating partners is an obstacle to the government taking the leading role.

- Donors need ensure that their expectations in terms of data availability and quality are realistic, sometimes accepting second-best options. The implementation of standardised sector information systems may make excessive demands on skills and resources.

- Dissemination of findings to national stakeholders and progress on systematic analysis are limited. The need for evidenced-based policy proposals and analysis of relevant data should be reinforced.

A number of recommendations for each of these key areas are outlined below:

- Donors should adopt a step-by-step approach, supporting, strengthening and coordinating existing activities that deliver reasonably reliable data, while encouraging the development of improved information systems for the medium and long term.

- PRS monitoring and assessment procedures should be integrated. Donors should encourage the evolution of a PRS policy matrix and commit to full alignment. In the interim, the focus on policy issues in existing assessment procedures will have to be maintained.

- Initially, data on budget allocations and expenditures, combined with a limited number of key output indicators, could provide the core of PRS implementation monitoring. The primary aim should be clarity and openness. Harmonisation between and within donor agencies should follow.

- A common-basket funding arrangement on information initiatives could be developed within a set of agreed rules. At a minimum, donors should consult on any proposed initiative relating to surveys, monitoring or information systems.

- The long-term aim should be the development of a statistics office with the capacity to oversee national and sectoral information systems. Community involvement in PRS monitoring can increase ownership, accountability and transparency but requires donor funding and clear terms of engagement.

- Initially, donor initiated and supported requests for data and information may remain vital. Encouraging governments to make findings accessible, and donor funding of local consultants and research institutions could also help.

Source: Lucas, H., and Zabel, M., 2005, 'The Role of Donors in PRS Monitoring Systems: Advantages and Disadvantages of the Different Forms of Involvement', Report prepared for the World Bank, HLSP Institute, London

Author: Henry Lucas and Monika Zabel , HLSP, http://www.hlsp.org/
⬤ Full text available online:
http://www.hlspinstitute.org/files/project/128177/WB_PRSP_note_061.pdf

---

2j. Civil Society's Perspective on their Involvement in PRSP Monitoring: Assessing Constraints and Potentials in Nicaragua

Author: T Braunschweig and B Stockli
Date: 2006
Size: 53 pages

What are the constraints and potentials facing civil society participation in Poverty Reduction Strategy (PRS) monitoring processes in Nicaragua? This Alliance Sud study reports on an ongoing initiative to develop an analytical tool to assess civil society's participation in PRS monitoring processes from their own perspective. It finds that basic conditions for a significant role for civil society organisations (CSOs) in the monitoring system in Nicaragua are not met.

Monitoring processes are central to the success of the PRS initiative. Civil society participation adds value to the process by providing input from deprived groups, increasing government accountability, strengthening ownership and broadening monitoring approaches. Flawed monitoring systems in most countries result in the potential contributions of CSOs being under-exploited. In Nicaragua, the credibility of the PRS process has been undermined by the perception that is was externally imposed. Lack of ownership and the failure of the government to genuinely incorporate CSOs in to related policy development have weakened their commitment to the PRS and its monitoring.

The lack of a culture of accountability in Nicaragua, low quality and inaccessible public information and range of broader factors relating to the political and social situation contribute to the limited monitoring role of CSOs. Key challenges include the following:

- Whilst there are many mechanisms for participation, CSOs perceive them to lack orientation and impact. Without effective means of pressure, CSOs have limited negotiating capacity with the government.

- Participatory monitoring exercises are hampered by people's general unfamiliarity with the concept of accountability and lack of requisite capacity or information.

- The government's failure to strive for a national consensus on poverty-reducing policies and lack of serious engagement with civil society is causing frustration and growing disinterest in the PRS monitoring process amongst CSOs.

- CSOs lack common strategies and sufficient collaborative links, weakening their ability to develop effective proposals and influence the monitoring process and policy at national level.

- Donors' funding bias towards and proximity to the government has debilitating effects on civil society's monitoring role. Support for CSOs is relatively limited and project-based, with competition between CSOs for funding undermining joint monitoring activities.

The focus for strengthening CSOs' monitoring role is on initiating and fostering processes that contribute to overcoming prevailing conditions in the long-term, rather than supporting specific monitoring activities. This draws on the potentials of CSOs - particularly at local level - and key fields of activity:

- CSOs have the capacity to strengthen human development at the local level to enhance participation in PRS monitoring in the long-term. Awareness raising and empowerment through the dissemination of information and engagement in the formulation and monitoring of local policies can contribute to building interest and ownership in the PRS.

- Investing in the research and analytical capacity of CSOs would increase their credibility as political actors and improve the effectiveness of lobbying and advocacy work.

- Effective contributions to the development of public policies would be improved through building on existing collaborative arrangements among CSOs, including umbrella organisations and networks.

- Recommendations to enhance donor assistance to CSOs include improving coordination through multi-donor funding and increased, and long-term core support to enable strengthening of institutional and government oversight capacities.

Source: Braunschweig, T. and Stöckli, B., 2006, 'Civil Society's Perspective on their Involvement in PRSP Monitoring: Assessing Constraints and Potentials in Nicaragua', Swiss Alliance of Development Organisations (Alliance Sud), Bern
Author: Thomas Braunschweig and Bruno Stöckli, Swiss Alliance of Development Organisations, http://www.alliancesud.ch/english/pagesnav/P.htm
Full text available online: http://www.alliancesud.ch/english/files/T_EgDyCa3.pdf

---

2k. A Guide to Designing a National Strategy for the Development of Statistics (NSDS)

Author: PARIS21
Date: 2004
Size: 37 pages

p. 11

There is increasing awareness of the need to strengthen statistical capacity to support the design, monitoring and evaluation of national development plans. National Strategies for the Development of Statistics (NSDS) are designed to achieve this goal. This guide prepared by the Partnership in Statistics for Development in the Twenty-first Century

(PARIS21) aims primarily to assist developing countries to design their NSDSs but will also be useful to development partners.

The NSDS should be integrated into national development policy processes and context, taking account of regional and international commitments. It should be developed in an inclusive way, incorporating results-based management principles and meet quality standards. It should be comprehensive and coherent and provide the basis for the sustainable development of statistics with quality. It ought to show where the statistical system is now, how it needs to be developed and how to accomplish this.

A good strategy should include mechanisms for consultation with the main stakeholders, an assessment of the current status of the system, an agreed vision, an identification of strategic actions, an action plan and mechanisms to monitor progress. Deciding on what agencies, data sets, and activities are to be covered by the NSDS in any one country will be a national decision. Nevertheless, the NSDS process should have as broad a coverage as possible:

- A broad NSDS process can promote greater coordination and cooperation in statistics and increase efficiency.

- It will help to bring the supply of statistics more in line with demand. Specific plans can then be prepared for improving data coverage, frequency, timeliness and other measures of quality.

- It can help develop the level of interest in, and support for, the development of the concept of National Statistics. It should build on experience in coordinating and improving the quality of National Statistics.

- The concept of the National Statistical System and the strategic plan should be demand-focused and user-friendly to maximize the added value of statistic outputs.

- Strategic planning should take into consideration use at different levels -national, international, and domestic regional.

The preparation of a good strategy will depend, crucially, on what mechanisms and processes are already in place. The following phases are recommended:

- Phase I: Launching the process (NSDS Design Road Map). This involves advocacy, sensitisation and dialogue with politicians, policy-makers, and decision-makers.

- Phase II: Assessment of the Current Status of the National Statistical System. This includes: collecting and analysing existing documentation; identifying user satisfaction, data needs and gaps; assessing outputs against quality criteria; assessing methodologies and the quality of statistics; assessing existing capacity to meet the needs and gaps; reviewing the legal and institutional framework, linkages, and coordination arrangements; and assessing organisational factors.

- Phase III: Developing the vision and identifying strategic options. This involves agreeing a mission and vision statements, agreeing on desired results and setting priorities and strategies to deliver the vision and results.

47

- Phase IV: Preparing the implementation plan. This requires a costed and time-bound action plan and a financial plan incorporating proposals for external assistance.
- Phase V: Implementation, monitoring and evaluation. Statistical systems must remain flexible. Strategic management should be a continuous process with mechanisms to monitor and evaluate progress, to review the strategy and to make modifications.

Source: PARIS21, 2004, 'A Guide to Designing a National Strategy for the Development of Statistics (NSDS)', The Partnership in Statistics for Development in the Twenty-first Century (PARIS21), Paris
Author: PARIS21 Secretariat, http://www.paris21.org
📄 Full text available online:
http://unstats.un.org/unsd/statcom/doc05/NSDS-Guide-English-PARIS21.pdf
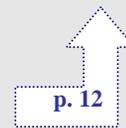

3a. When Will We Ever Learn? Improving Lives Through Impact Evaluation

Author: Center for Global Development
Date: 2006
Size: 95 pages

Despite decades of investment in social development programmes, we still know relatively little about their net impact. So why are rigorous social development impact evaluations relatively rare? This paper from the Center for Global Development (CGD) aims to address this question and provide recommendations for more and better evidence for policymaking and programme planning. A new, collective approach is needed, in which developing country governments, bilateral and multilateral development agencies, foundations and NGOs work together to define an agenda of enduring questions and fund the design and implementation of rigorous impact evaluations in key sectors.

Although significant resources are devoted to monitoring of programme outputs or process evaluations, there is insufficient investment in impact evaluations. Reviews of evidence that are conducted in the course of policy or programme design routinely reveal that little is known about "what works" and decisions are made on the basis of anecdotal or other weak evidence. Of those studies that have been completed, many are methodologically flawed so that net impact cannot be estimated in a valid way.

Three basic incentive problems need to be overcome to generate more and better impact evaluations:

- A portion of the knowledge generated through impact evaluation is a public good. Such broad benefits are amplified greatly when the same type of programme is evaluated in multiple contexts, and addresses enduring questions. However, the cost-benefit calculation for such an evaluation might not include these benefits, making it look like the impact evaluation is not worthwhile.

48

- The rewards for institutions, and for individuals within them, come from "doing" not from "building evidence" or "learning". Thus, it is extremely difficult to protect the funding for good evaluation, or to delay the initiation of a project to design the evaluation and conduct a baseline study. The opportunity costs are seen as too great. Consequently, resources that might be devoted to rigorous evaluation are used instead for project implementation.

- There are disincentives to finding out the truth. If policymakers and programme managers believe that future funding depends directly on achieving a high level of success, the temptation will be to concentrate on producing and disseminating "success stories".

The solution to overcoming these incentive problems lies both in strengthening capacity within organisations and in creating a collective approach. Specifically:

- Organisations should commit individually to strengthening monitoring and evaluation systems, dedicating resources to impact evaluations and building capacity in developing countries.

- These efforts by individual organisations should be coupled with collaborative - collective action among actors to coordinate and fund rigorous impact evaluation.

- A committee, standards-based network, secretariat, or other organisation is needed as a focal point to lead a collective action initiative. This "council" should include any developing country governments, development agencies, NGOs, foundations, and other public and private entities willing to engage actively in improving the quality and increasing the quantity of impact evaluations.

- The core functions of the council would be: to establish quality standards; administer a review process for evaluation designs; identify priority topics; and provide grants. Other functions would include: organising and disseminating information; building capacity to produce, interpret, and use knowledge; creating a directory of researchers; and undertaking communication activities.

Source: Center for Global Development, 2006, 'When Will We Ever Learn? Improving Lives Through Impact Evaluation', Evaluation Gap Working Group, Centre for Global Development, New York
Author: Center for Global Development, http://www.cgdev.org
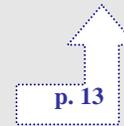 Full text available online:
http://www.cgdev.org/files/7973_file_WillWeEverLearn.pdf

3b. Use of Randomization in the Evaluation of Development Effectiveness

Author: E Duflo and M Kremer
Date: 2003
Size: 37 pages

Just as randomised pharmaceutical trials revolutionised medicine in the 20th Century, randomised evaluations could revolutionise social policy in the 21st. This paper, prepared for a 2003 World Bank Operations Evaluation Department (OED) conference, draws on evaluations of educational programmes. It argues that there is an imbalance in evaluation methodology and recommends greater use of randomised evaluations. As credible impact evaluations, these could offer valuable guidance to international organisations, governments, donors and NGOs in the search for successful programmes.

Impact evaluation asks how those who participated in a programme would have fared without it. Comparing the same individual over time is problematic, since other things may have changed. Thus, it is critical to establish a comparison group. But such groups can be affected by pre-existing differences (selection bias) or by the programme itself. One way of overcoming these issues is to randomly select treatment and comparison groups from a potential population of participants. Other techniques to control for bias include propensity score matching, difference-in-difference estimates and regression discontinuity design.

The following insights are drawn from randomised evaluations of educational programmes:

- Progresa, a welfare programme in Mexico, transformed a budgetary constraint into an opportunity for evaluation by randomly selecting the order in which participant communities would be phased in. Evaluation showed improvements in health and education and led to programme expansion.

- Evaluating several programmes in western Kenya enabled comparison of the cost-effectiveness of approaches to increasing school participation.

- Randomised evaluation of textbook and flipchart provision in Kenyan schools suggested that retrospective analysis had overestimated effectiveness.

- An Indian NGO placed a second teacher in non-formal education centres to overcome staff and pupil absenteeism. Evaluation informed the decision not to scale-up the programme on the grounds that benefits failed to outweigh costs.

- Some parent-run school committees in Kenya provide gifts to teachers whose students perform well. Analysis indicates that teachers respond with efforts to manipulate test scores rather than stimulating long-term learning.

- A Colombian programme allocated private school vouchers through a lottery system, renewable depending on academic performance. The benefits of the programme exceeded the costs, with winners more likely to graduate and score highly on exams.

50

Randomised evaluations are often feasible and can shed light on issues outside their specific focus. International organisations have an important role to play in supporting credible evaluations:

- Governments are not the only outlets through which randomised evaluations can be organised. NGOs are well-suited to conducting them, but require technical assistance and outside financing.

- Conducting a series of evaluations in the same area can reduce costs and enhance comparability. Once staff are trained, they can work on multiple projects and share data.

- Not all programmes are suitable for randomised evaluation: programmes targeted to individuals or communities are stronger candidates than programmes which operate only at the central level, such as central bank independence. Even then, not all projects need impact evaluation and proposals should be carefully considered before money is spent.

- There is scope to set up a specialised unit within the institutional environment of international organisations to encourage, conduct and finance rigorous impact evaluations. Such a unit could publish a database of impact evaluations to help alleviate problems of publication bias.

Source: Duflo, E., and Kremer, M., 2003, 'Use of Randomization in the Evaluation of Development Effectiveness', Paper prepared for the World Bank Operations Evaluation Department (OED) Conference on Evaluation and Development Effectiveness 15-16 July, 2003, Massachusetts Institute of Technology , Cambridge, Mass
Author: Esther Duflo and Michael Kremer, Department of Economics, Massachusetts Institute of Technology (MIT), http://econ-www.mit.edu/
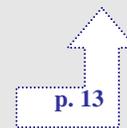Full text available online:
http://econ-www.mit.edu/facultyload_pdf.php?id=759

3c. Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate Progresa

Author: O Attanasio et al
Date: 2005
Size: 38 pages

What impact have monetary incentives had on education choices in rural Mexico? How can the design of educational interventions aimed at improving educational participation be improved? This paper from the Institute for Fiscal Studies (IFS) analyses the education component of the Mexican government's welfare programme, Progresa, which aims to reduce rural poverty. It argues that increasing the grant for secondary school children while eliminating it at the primary age would strengthen Progresa's impact.

Progresa grants aim to keep children in school. They start at age 6 and increase until age 17 and are conditional on enrolment and attendance. The evaluation of Progresa is helped by the fact that data has been collected from the outset. This paper combines the information provided by a randomized allocation of Progresa across localities with a simple structural model of education choices. This gives a better picture of Progresa's specific mechanisms and supports identification of possible improvements.

Since the structural model is based on boys, results are only reported for them:

- Progresa has, on average, a 3% effect on enrolment of boys aged 6 - 17. This effect is greater for older boys (around 7.5%) and virtually zero for boys under 9. Almost all children go to school below grade six (11 years) so the early part of the grant is not truly conditional.

- Belonging to a household with less educated parents leads to lower attendance. Children from poor households have, on average, lower levels of schooling.

- Pre-existing education is another determinant of choice: increased levels of education have a positive effect on further participation. Children living in villages with greater availability of schools are better educated.

- Control villages were affected by the fact that it was known the programme would be implemented there in the future.

- Cost variables had the expected effect: The higher the cost of attending school, in terms of distance or financial costs, the less likely children are to attend.

Some implications of and improvements to Progresa are suggested, as are contributions to the evaluation literature:

- Conditional grants are a relatively effective means of increasing the enrolment of children at the end of their primary education.

- The impact on school participation would almost double by restructuring the grant to target older children. There would be no negative effect on the school participation of primary age children.

- One alternative policy - though not equivalent in terms of cost and other benefits - would be a school building programme. By reducing the distance of secondary schools to no more than 3km, participation would increase by 6% at age 15.

- Without a structural model of education choices, it is impossible to evaluate the effect of the programme and possible changes to its structure. The randomised component of the data enables flexibility to be built into the model.

3d. Conducting Quality Impact Evaluations Under Budget, Time and Data Constraints

Author: M Bamberger
Date: 2006
Size: 35 pages

How do cost, time and data constraints affect the validity of evaluation approaches and conclusions? What are acceptable compromises and what are the minimum methodological requirements for a study to be considered a quality impact evaluation? This booklet from the World Bank provides advice for conducting impact evaluations and selecting the most rigorous methods available within the constraints faced. It clarifies the nature of trade-offs between evaluation rigour and budgets, time and data and provides suggestions for reducing costs and increasing rigour.

A quality impact evaluation must: (a) define and measure project inputs, implementation processes, outputs, intended outcomes and impacts; (b) develop a sound counterfactual; (c) determine whether a project has contributed to the intended impacts and benefited the target population; (d) assess the distribution of benefits among the target population; (e) identify factors influencing the magnitude and distribution of the impacts; and (f) assess the sustainability of impacts.

The challenge of conducting evaluations under budget, time and data constraints is outlined below:

- The conventional evaluation approach uses pre- and post-intervention project and control group comparisons. Simplifying evaluation design can reduce costs and time, but involves trade-offs between evaluation quality and the validity of conclusions.

- Options for selecting comparison groups include matching areas, individuals or households on observables, pipeline sampling, regression discontinuity designs and propensity score matching. Problems include self-selection and difficulties finding close matches.

- Secondary data can be used to reduce costs and save time and to reconstruct baseline data and comparison groups, thus strengthening counterfactuals.

- Post-intervention only evaluations suffer from a lack of baseline data. Baseline data may be reconstructed using project and other records, recall, Participatory Rapid Assessment (PRA) or key informants. There should be systematic comparison of the information obtained.

- Fewer interviews, cost-sharing, shorter and simpler survey instruments and electronic and community-level data collection can reduce data collection costs. Using nurses, teachers, students or self-administered questionnaires can reduce interview costs.

Whilst trade-offs between evaluation rigour and constrained resources are inevitable, there are a number of ways of minimising the effects of constraints on the validity of evaluation conclusions:

- Considering the cost implications of addressing selection bias and instrumental variables can strengthen overall evaluation design quality. Addressing selection bias in project design is often more economical than during post-implementation.

- Developing programme theory models can help identify areas to concentrate resources. Peer review can help assess threats to validity when addressing constraints. Preparatory studies completed before the arrival of foreign consultants help to address time constraints.

- When using smaller sample sizes, cost-effective mixed-method approaches can strengthen validity by providing multiple estimates of key indicators. When reducing the sample size, statistical power analysis can ensure the proposed samples are large enough for the analysis.

- Reconstructing baseline conditions and devoting sufficient time and resources to developing programme theory can strengthen theoretical frameworks and the validity of counterfactuals. Rapid assessment studies are cost-effective for developing programme theory models.

- Rapid assessment methods to identify similarities and differences between project and comparison groups as well as contextual analysis of local factors can strengthen the generalisabilty of conclusions.

- An evaluability assessment should be conducted to determine whether a quality impact assessment is possible. If not, the evaluation resources, time frame and objectives should be revised, or the evaluation cancelled.

Source: Bamberger, M., 2006, 'Conducting Quality Impact Evaluations Under Budget, Time and Data Constraints', Independent Evaluation Group, World Bank, Washington, D.C.
Author: Michael Bamberger, World Bank Independent Evaluation Group, http://www.worldbank.org/ieg/ecd/
 Full text available online:
http://lnweb18.worldbank.org/oed/oeddoclib.nsf/DocUNIDViewForJavaSearch/757A5C C0BAE22558852571770059D89C/$file/conduct_qual_impact.pdf

3e. RealWorld Evaluation: Working Under Budget, Time, Data and Political Constraints: An Overview

Author: M Bamberger et al
Date: 2006
Size: 75 pages

How can the principles of optimal evaluation design be applied under real world conditions with budget, time, data and political constraints? This paper, adapted from chapter 16 of RealWorld Evaluation: Working under Budget, Time, Data and Political Constraints provides an overview of the RealWorld Evaluation (RWE) approach. It addresses constraints through practical suggestions applicable to both developing and developed world research. Understanding the aims and purpose of the evaluation, as well as the local context, is critical.

RWE approaches can be applied at each stage of the design and implementation of a typical evaluation. This requires careful planning and design:

- Evaluators need a clear understanding of clients' and stakeholders' needs and expectations and the political environment of the evaluation. It is important to consider the dynamics of power and relationships of the key players in the project. Philosophical or methodological orientations must also be recognised.

- Defining the explicit or implicit theory or logic model of the project to be evaluated is critical. This model should include the following phases: design, inputs, implementation process, outputs, outcomes, impacts and sustainability.

- The seven most widely used RWE designs range from longitudinal quasi-experimental, to pre-post-test without control, to simple end-of-project assessment without baseline or comparison group. Less rigourous designs require techniques to cope with missing data.

- Further considerations include streamlining the evaluation design, identifying what analysis and comparisons are critical and assessing threats to validity and adequacy of different designs.

- Evaluators should select the tools best suited to the needs of the client and the nature of the evaluation. In most situations the strongest and most robust evaluation design will probably combine both quantitative and qualitative approaches.

A number of steps can be taken to increase the effectiveness of an evaluation in real world situations:

- Evaluation designs can be strengthened by: basing the evaluation design on a programme theory model; complementing the results-based evaluation design with a process evaluation; incorporating contextual analysis; reconstructing baseline conditions to assess changes in key indicators over the life of the project; and triangulation, including obtaining the perspectives of a range of stakeholders.

- The evaluation team should include people with different experiences, skill sets and perspectives. External consultants should be limited to essential areas. At least one team member needs to be a content specialist. Data collectors can be resourced creatively to save costs.

- Data collection plans need to be simplified. Preparatory studies and reliable secondary data can save time and expense. Simple ways to collect data on sensitive topics and difficult to reach populations need to be found.

- The analysis and report should focus on answering key questions. Limitations in the validity of conclusions should be clearly acknowledged, with reference to the RealWorld Evaluation integrated checklist for assessing the adequacy and validity of quantitative, qualitative and mix-method designs.

- Reports should be succinct, of direct practical utility to different stakeholders and communicated in appropriate ways to a range of audiences.

- A follow-up action plan should be developed with the client to ensure the findings and recommendations are well used.

Source: Bamberger, M., Rugh, J., and Mabry, L., 2006, 'RealWorld Evaluation: Working under Budget, Time, Data and Political Constraints', Sage Publications, California
Author: Michael Bamberger, Jim Rugh and Linda Mabry
 Full text available online:
http://www.realworldevaluation.org/uploads/RealWorld_Evaluation_-_Summary_Chapter.doc

---

3f. Impact Evaluation: The Experience of the Independent Evaluation Group of the World Bank

Author: H White
Date: 2006
Size: 58 pages

Aid spending is increasingly dependent on proof that interventions are contributing to the attainment of Millennium Development Goals (MDGs). Yet there is still debate over the definition of impact evaluation and how it should be carried out. This paper draws on the experience of the Independent Evaluation Group (IEG) of the World Bank. It defines impact evaluation as a 'counterfactual analysis of the impact of an intervention on final welfare outcomes' and recommends a theory-based approach. Two sources of bias are highlighted: contamination and self-selection bias.

Debates over impact evaluation reflect wider discussion of qualitative and quantitative research methods. There is a balance to be struck between rigour and policy relevance. A theory-based approach looks at how a programme is working, not just if it is working. This requires quantitative and qualitative data. The challenge is the search for the counterfactual: a comparison between what happened and what would have happened

56

without intervention. This means selecting a comparison group that is like the treatment group in every way but has not been subject to the intervention.

Two problems can lead to misleading results:

- Contamination (or contagion): Since the comparison group has to be similar to the treatment group, it may be indirectly affected by the intervention. This can be through geographical spillover or similar interventions elsewhere.
- Self-selection bias: Since beneficiaries chosen through self-selection are not a random sample of the population, random selection should not be used to construct the comparison group.

Four case studies illustrate IEG's experience of impact evaluation:

- Improving the quantity and quality of basic education in Ghana: While showing improvement in enrolment and outcomes, evaluation highlighted a lack of physical resources and problems with teacher morale.
- Meeting health MDGs in Bangladesh: Evaluation confirmed publicly-provided services with external support to be cost-effective in improving health outcomes. Local evidence was shown to be important in resource allocation.
- The Bangladesh Integrated Nutrition Project (BINP): Evaluation showed only low-level reduction in malnutrition. Key decision-makers had been neglected in nutritional counselling, as had other factors constraining nutritional practices.
- Agricultural extension services in Kenya: Evaluation revealed greater impact in poorer areas, though delivery had been concentrated in more productive areas. However, programme impact was limited because recommended changes in extension practices were only partially implemented, and extension advice was outdated.

Recommendations relate to overcoming problems of bias and producing policy-relevant results:

- A theory-based approach gives a picture of the whole causal chain from inputs to impacts, where some evaluations only look at final outcomes. This means mapping out how inputs are expected to achieve intended outcomes.
- To avoid contamination, data on interventions in the comparison group should be collected. A theory-based design is better able to deal with different types and levels of intervention.
- A variety of experimental and quasi-experimental methods can be used to address issues of selection bias
- Time for survey-design and testing should not be underestimated and will help teams write well-contextualised reports. Staff collecting data should be fully trained and avoid leading questions.
- While establishing the counterfactual is crucial, the factual should not be overlooked. Using data from the treatment group to establish a picture of what has

actually happened is an essential part of impact evaluation.

Source: White, H., 2006, 'Impact Evaluation: The Experience of the Independent Evaluation Group of the World Bank', World Bank, Washington, D.C.
Author: Howard White , World Bank, http://www.worldbank.org/oed
 Full text available online:
http://mpra.ub.uni-muenchen.de/1111/01/MPRA_paper_1111.pdf

3g. Mixed-Method Evaluation

Author: M Bamberger et al
Date: 2006
Size: 20 pages

p. 15

Quantitative and qualitative methods of research each have strengths and weaknesses when applied in isolation. However, combining the two approaches through mixed-method evaluation is gaining wider acceptance among social science researchers as a way of conducting more comprehensive and robust analysis. This chapter from RealWorld Evaluation: Working Under Budget, Time, Data and Political Constraints discusses the most appropriate contexts and strategies for using a mixed-method approach. It argues that mixed-method evaluation is a flexible and practical technique which can be used at any stage of an evaluation. Nevertheless, a fully integrated approach requires extensive planning and deliberation to ensure that the most appropriate combination of methods is chosen and successfully implemented.

Mixed-method evaluation combines the detailed insights and holistic understanding obtained from qualitative research with the ability to generalise to a wider population offered by quantitative data collection. Thus, it allows for a more comprehensive analysis. Mixed-method designs can be employed to strengthen validity, fine-tune sampling and instrumentation, extend the coverage of findings, conduct multi-level analysis and generate new and diverse insights:

- Mixed-method evaluation designs can give equal weight to qualitative and quantitative approaches, or allow one method to figure more prominently than the other.

- The different methods can be used concurrently or sequentially. Sequential designs are often easier to organise, although they may require more time if the second phase of research cannot begin until the first is completed. Concurrent designs can be logistically more difficult to manage, particularly if evaluation teams lack extensive experience of coordinating quantitative and qualitative methods simultaneously.

- Mixed-methods can be fully integrated throughout the evaluation process or used at any individual stage of the evaluation. However, in practice, mixed methods are generally applied at only one or two stages, be it hypothesis formulation, data collection, analysis and follow-up, or presentation and dissemination of findings.

58

- Mixed-methods permit multi-level analysis through the comparison of findings from data collected at the level of the individual household, group, organisation or community

In practice, undertaking a mixed-methods design has important implications for the planning and implementation of the evaluation. The benefits of mixed-method approaches vary depending on the specific weighting and application of the combination selected. The objectives of the project, along with resource constraints, will dictate the nature of the most useful combination of methods:

- Concurrent designs may save time overall, but require a large amount of planning and organisation compared to sequential approaches.

- While some evaluators refer to the use of a mixed-method design when they have only included additional data collection methods to a dominantly quantitative or qualitative approach, this is a misunderstanding of the approach. A mixed-method approach requires an integrated strategy in which the strengths and limitations of quantitative and qualitative methods are recognised and the evaluation is designed to take advantage of the complementarities between the different methods.

- A fully-integrated approach requires ensuring an interdisciplinary perspective at all stages of the research, including: composition of the research team, designing the evaluation framework, data collection, and analysis and follow-up.

- Fully-integrated approaches will generally require more time. Where research teams involve professionals from multiple disciplines, it is essential to invest additional time during the planning stage of the evaluation for building relations and common understandings among team members.

- One ongoing challenge is to develop sampling procedures that ensure subjects and cases for quantitative and qualitative data collection are drawn from the same universe. Frequently, it is difficult to know whether differences in the findings reflect real differences in the information obtained or are partly due to the different selection procedures used in each part of the study.

Source: Bamberger, M., Rugh, J., and Mabry, L., 2006, 'Mixed-Method Evaluation', Chapter 13 in RealWorld Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage Publications, California
Author: Michael Bamberger, Jim Rugh and Linda Mabry
 Full text available via document delivery:
http://www2.ids.ac.uk/docdel/grc2/grcdel2an.cfm?rn=258497

3h. Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners

Author: J L Baker
Date: 2000
Size: 217 pages

There is broad evidence that developmental assistance benefits the poor, but how can we tell if specific projects are working? Have resources been spent effectively? What would have happened without intervention? This comprehensive handbook from Directions in Development seeks to provide tools for evaluating project impact. It advises that effective evaluations require financial and political support, early and careful planning, participation of stakeholders, a mix of methodologies and communication between team members.

To be effective, an evaluation must estimate the counterfactual: what would have happened without intervention? This is at the core of evaluation design. In experimental design (randomisation), intervention is allocated randomly, thus creating treatment and control groups statistically equivalent to one another. Quasi-experimental design uses non-random methods to generate a comparison group similar to the treatment group. This is normally done after the intervention has taken place, so statistical controls are applied to address differences between the groups. Qualitative techniques focus on understanding processes, behaviours and conditions as observed by those being studied.

There are a number of key steps in designing and implementing impact evaluation:

- Determining whether or not to carry out an evaluation; assessing the costs and benefits; and considering other options. If an evaluation is to go ahead, objectives have to be clarified and availability of data assessed.

- The choice of methodologies will depend on the evaluation question, timing, budget constraints and implementation capacity. Evaluation design should consider how quantitative and qualitative methods can be integrated to complement each other.

- Impact evaluation requires a range of skills. Team members should be identified, responsibilities agreed and communication mechanisms established.

- Planning data collection should cover: sample design and selection; instrument development; staffing and training; pilot testing; data collection; and data management and access.

- Data collection should be followed by analysis, which often takes at least a year. Content analysis reviews quantitative data from interviews and observation, whereas case analysis is used for qualitative data.

- Findings should be written up and discussed with policymakers and other stakeholders before being incorporated into project design.

The following recommendations are drawn from 'good practice' impact evaluations:

- Early and careful planning of the evaluation design: Ensure the right information is collected and allow for the use of results in mid-course adjustments.

- Approaches to evaluation when there is no baseline: Where it is not possible to generate control data during the evaluation, this can be constructed using matching methods and existing data.

- Dealing with constraints on developing good controls: Where randomisation or experimental controls are not politically feasible, evaluations can make use of pilot projects in restricted areas.

- Combining methods: Using qualitative and quantitative techniques provides both the quantifiable impact of a project and an explanation of how processes came about.

- Exploiting existing data sources: Drawing on sources such as national surveys, censuses or administrative records reduces the need for costly data collection.

- Costs and financing: Data collection and consultants tend to be expensive components. While many countries assume the bulk of costs, impact evaluation often requires substantial outside resources.

Source: Baker, J., 2000, 'Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners', World Bank, Washington
Author: Judy L. Baker, World Bank, www.worldbank.org
Full text available online:
http://siteresources.worldbank.org/INTISPMA/Resources/handbook.pdf

---

3i. Impact Evaluation and the Project Cycle

Author: World Bank
Date: 2006
Size: 26 pages

p. 15

The goal of an impact evaluation (IE) is to attribute impacts to a project using a comparison group to measure what would have happened to the project beneficiaries had it not taken place. The process of identifying this group, collecting the required data and conducting the relevant analysis requires careful planning. This paper from the World Bank provides practical guidance on designing and executing IE. It includes some illustrative costs and ideas for increasing government buy-in to the process.

A good IE provides the basis for sound policy making, revealing the size of a project's impact and the identity of its beneficiaries. In addition to providing hard evidence with which to weigh and justify policy priorities, IE can also be used as a managing-by-results tool. As the evaluation progresses side by side with the project, it can test project features to modify design and improve effectiveness over time. Among other things, IE can help policy makers examine the effect of a pilot, compare different delivery modes, and examine the impact of the project for different populations. Overall, IE helps to build and

61

sustain national capacities for evidence-based policy making by allowing us to learn which projects work in which contexts for the next generation of policies.

Whilst there is no standard approach to conducting an IE, there are identifiable periods in the life of a project: a) from identification of the project through to the writing of the PCN; b) from preparation through to appraisal; and c) from appraisal through to completion. To complete a successful evaluation, each period may be divided into six sections, representing different types of activity:

- Evaluation activities: the core issues surrounding the evaluation design and implementation
- Building and maintaining constituencies: dialogue with the government and World Bank management
- Project considerations: logistical considerations relating to the evaluation
- Data: issues related to building the data needed for the evaluation
- Financial resources: considerations for securing funding
- Support: resources available to project leaders to assist in the evaluation process

Each evaluation must be tailored to the specific project, country and institutional context, and to the actors involved. In the planning of any IE, however, there are some overarching themes to address:

- Integrating the IE with the project – The members of the IE team must become intimately familiar with the aims, context, possible design options and execution of the project. Equally, the project leaders and clients must engage fully with the process to gain maximum positive results.
- Relevance – Evaluations must answer questions when feedback can still be incorporated into policy-making, such as project mid-term reviews, budget discussions, PRSP preparations or progress reports.
- Government ownership – This is central to success of the project in order to identify relevant policy questions, ensure integrity/relevance of the evaluation, and incorporate results into future policy-making and resource allocation. Early collaboration between government, project leaders and the IE team is essential.
- Flexibility and adaptability – The evaluation must adapt to context, paying close attention to the political environment, in order to capitalise on planned changes and respond to fluctuations in support and opposition.
- Timing – An adequate results framework is required to ameliorate potential tension between evaluation objectives and the project timeframe. Some long-term outcomes may require a financial strategy that allows for evaluation beyond project completion.

Source: The World Bank, 2006, 'Impact Evaluation and the Project Cycle', PREM Poverty Reduction Group, World Bank, Washington, D.C.

Author: Tara Bedi, Sumeet Bhatti, Xavier Gine, Emanuela Galasso, Markus Goldstein and Arianna Legovini, World Bank , www.worldbank.org
🔘 Full text available online:
http://siteresources.worldbank.org/INTISPMA/Resources/383704-1146752240884/doing_ie_series_01.pdf

---

4a. Ten Steps to a Results-based Monitoring and Evaluation System

Author: J Kusek and R Rist
Date: 2004
Size: 268 pages

p. 16

Governments and organisations face increasing internal and external pressures to demonstrate accountability, transparency and results. Results-based monitoring and evaluation (M&E) systems are a powerful public management tool to achieve these objectives. This handbook from the World Bank presents a ten-step model that provides extensive detail on building, maintaining and sustaining a results-based M&E system.

Results-based M&E systems can help build and foster political and financial support for policies, programmes and projects and can help governments build a solid knowledge base. They can also produce major changes in the way governments and organisations operate, leading to improved performance, accountability, transparency, learning, and knowledge. Results-based M&E systems should be considered a work in progress. Continuous attention, resources, and political commitment are needed to ensure their viability and sustainability. Building the cultural shift necessary to move an organisation toward a results orientation takes time, commitment and political will.

The ten steps to building, maintaining and sustaining a results-based M&E system are outlined below:

- A readiness assessment should be conducted to determine whether prerequisites for a results-based M&E system are in place. It should review incentives and capacity for an M&E system and roles, responsibilities and structures for assessing government performance.

- Outcomes to monitor and evaluate should be agreed through a participatory process identifying stakeholders' concerns and formulating them as outcome statements. Outcomes should be disaggregated and a plan developed to assess how they will be achieved.

- Key performance indicators to monitor outcomes should be selected through a participatory process considering stakeholder interests and specific needs. Indicators should be clear, relevant, economical, adequate and monitorable.

- Baseline data on indicators should be established as a guide by which to monitor future performance. Important issues when setting baselines and gathering data on indicators include the sources, collection, analysis, reporting and use of data.

63

- Performance targets should be selected to identify expected and desired project, programme or policy results. Factors to consider include baselines, available resources, time frames and political concerns. A participatory process with stakeholders and partners is key.

- Monitoring for results includes both implementation and results monitoring as well as forming partnerships to attain shared outcomes. Monitoring systems need ownership, management, maintenance and credibility. Data collection needs reliability, validity and timeliness.

- Evaluation provides information on strategy, operations and learning. Different types of evaluation answer different questions. Features of quality evaluations include impartiality, usefulness, technical adequacy, stakeholder involvement, value for money and feedback.

- Reports on the findings of M&E systems can be used to gain support and explore and investigate. Reports should consider the requirements of the target audience and present data clearly.

- Findings of results-based M&E systems can also be used to improve performance and demonstrate accountability and transparency. Benefits of using findings include continuous feedback and organisational and institutional knowledge and learning.

- Good results-based M&E systems must be used in order to be sustainable. Critical components of sustaining M&E systems include demand, clear roles and responsibilities, trustworthy and credible information, accountability, capacity and incentives.

Source: Kusek, J., and Rist, R., 2004, 'Ten Steps to a Results-based Monitoring and Evaluation System', World Bank, Washington, D.C.
Author: Jody Zall Kusek and Ray Rist, World Bank , www.worldbank.org
Full text available online: http://www.oecd.org/dataoecd/23/27/35281194.pdf

---

4b. Guidance on Evaluation and Review for DFID Staff

Author: Department for International Development
Date: 2005
Size: 89 pages (1.3 MB)

p. 16

Good evaluation practice depends on a solid partnership between those commissioning and managing evaluation studies, and the consultants undertaking the work and producing reports. This guide from the UK Department for International Development (DFID) aims to improve the quality of decentralised evaluation. It outlines the steps for designing, managing, reporting on and responding to an evaluation.

Assessments of development assistance should describe what has happened and why, using reliable and transparent methods of observation and analysis. Within DFID, this

64

takes the form of impartial, independent and in-depth evaluations leading to an independent report. Internal evaluation activities, taking place during the lifetime of the project are referred to as reviews.

Evaluations help development partners to learn from their experiences and provide accountability for the use of resources to Parliament and taxpayers. Assessing the impact of global and national development activities requires ever more complex, multi-stakeholder, thematic evaluations. Some examples include:

- evaluations categorised by when they take place, the processes used and by where they focus: for example, annual reviews or participatory evaluations.

- project evaluations which focus on a pre-defined cause-effect chain, leading from project inputs to outputs and fulfilment of objectives.

- programme evaluations which focus on questions of institutional performance, processes, changes and interrelationships, as well as the development impact of the programme area.

- pvaluations of aid instruments which focus on the efficiency, effectiveness and impact of an aid instrument and its management between development partners.

- country development evaluations where the partner country is in the driving seat, reflecting on the value added by the combined efforts of the international development community.

Well-planned and -managed evaluation exercises can improve the quality of future development activities. There are number of factors to consider when using evaluations and sharing lessons:

- The evaluation manager should think about the purpose of the evaluation and who the main users of findings may be. Involving potential users of the findings in evaluation design is the best way to ensure interest in the results.

- The process of conducting the evaluation is useful for sharing lessons. Local officials wishing to learn more about the programme and its implementation may accompany the evaluation team.

- Lessons can be shared beyond those immediately involved in the evaluation through planned dissemination activities. These include circulating reports to interested parties, arranging dissemination events, and organising press releases.

- The executive summary is a key part of the report for stakeholders and the public. It should be short and easily digestible and preferably translated into the language of the host country.

- It is sensible to ensure that a management response to the report is prepared before publication. This can be released with the report to show changes being implemented as a result of the recommendations.

Source: DFID, 2005, 'Guidance on Evaluation and Review for DFID Staff', Evaluation Department, Department for International Development, London

Author: Department for International Development (DFID), http://www.dfid.gov.uk

🔴 Full text available online:
http://www.dfid.gov.uk/aboutDFID/performance/files/guidance-evaluation.pdf

---

4c. Handbook on Monitoring and Evaluating for Results

Author: United Nations Development Programme Evaluation Office
Date: 2002
Size: 119 pages

Since 1999, the United Nations Development Programme (UNDP) has placed greater emphasis on results in its work to eliminate poverty. That shift has led to new demands on Monitoring and Evaluation (M&E) in country offices and programme units. This handbook outlines an M&E framework for use by UNDP staff and partners that promotes learning and performance measurement.

The goal of the UNDP's reform programme is to demonstrate how and where the organisation is making a measurable contribution to reducing poverty. It is centred on Results-based Management (RBM), a methodology in which performance is systematically measured and improved and resources are used to enhance effectiveness. The practical guidance offered in the handbook is intended to support country offices in aligning their M&E systems with RBM methodology. It focuses on tracking and measuring the performance of UNDP interventions and strategies, and their contributions to outcomes. The proposed framework is meant to foster an organisational culture of learning, transparency and accountability, and aims to simplify policies and procedures. The publication is expected to boost capacity for results-oriented development within UNDP and its national partners.

Monitoring is an ongoing function that aims to provide early indications of progress or lack of it towards achieving results. Evaluation, on the other hand, is an exercise that aims to systematically and objectively assess progress towards and/or the achievement of an outcome. Key features of the new UNDP approach include the following:

- The focus of monitoring is now on the outcomes of interventions, rather than project outputs as before.
- Outcome monitoring is periodic so that change can be perceived over time and compared against baseline data and outcome indicators.
- Outcome monitoring focuses on projects, programmes, partnerships, 'soft' assistance (policy advice, advocacy etc) and implementation strategies.
- Outcome evaluations are used to assess how and why the outcomes of a set of related activities may or may not be achieved in a country, and the role played by UNDP.

- They can help clarify underlying factors, reveal unexpected consequences, recommend actions to improve performance and generate lessons learned. They may take place at different times in a programme cycle.
- Programme managers are encouraged to involve partners actively in M&E.

Knowledge gained from M&E is at the heart of the UNDP's organisational learning process. An M&E framework that produces knowledge, promotes learning and guides action is an important means of capacity development. To achieve this:

- evaluative evidence should have 'real-time' capability, allowing it to be verified quickly and results given to stakeholders when most useful.
- it is important to learn what works in terms of outcome relevance, partnership strategy, output design and indicators, and to feed this back into ongoing and future work.
- evaluations should be seen as part of an exercise allowing stakeholders to participate in generating and applying knowledge.
- staff should record and share lessons learned with others, using evaluations as an opportunity to bring development partners together and spread learning beyond UNDP.
- country offices should plan and organise evaluations so that they cover the most crucial outcomes, are timely and generate sufficient information on lessons learned.
- early signs of potential problems detected by monitoring must be acted on. This serves immediate needs and may provide feedback for future programming.

Source: United Nations Development Programme Evaluation Office, 2002, Handbook on Monitoring and Evaluating for Results, UNDP, New York
Author: UNDP, www.undp.org
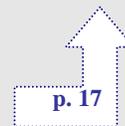
Full text available online:
http://www.un.org.pk/undp/prm/Rules%20&%20Guidelines/handbook-on-monitoring-and-evaluating-for-results-17april-02.pdf

---

4d. Principles for Evaluation of Development Assistance

Author: OECD-DAC
Date: 1991
Size: 11 pages

Aid evaluation plays an essential role in efforts to enhance the quality of development cooperation. This paper from the OECD's Development Assistance Committee presents a set of principles on the most important requirements of the evaluation process. Development assistance is a cooperative partnership between donors and recipients. Both must take an interest in evaluation to improve the use of resources through learning and to ensure accountability to political authorities and the public.

An evaluation is defined as a systematic and objective assessment of a project, programme or policy and its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives and the efficiency, effectiveness, impact and sustainability of the outcomes. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of recipients and donors.

The principles set out below provide general guidance on the role of aid evaluation in the aid management process, with the following central messages:

- Aid agencies should have an evaluation policy with clearly established guidelines and methods. The policy should have a clear definition of its role and responsibilities and its place within the institutional aid structure.

- The evaluation process should be impartial and independent from the policy-making process and the delivery management of development assistance.

- The evaluation process must be as open as possible with results made widely available.

- To be useful, evaluations must be used. Feedback to policy-makers and operation staff is essential.

- Partnership between donors and recipients is essential for the evaluation process. This is an important aspect of recipient institution building and of aid coordination and may reduce administrative burdens on recipients.

- Aid evaluation and its requirements must be an integral part of aid planning from the start. Clearly identifying the objectives of an aid activity is an essential prerequisite for objective evaluation.

In light of these principles, aid agencies need to ensure that they have a clear and defined set of guidelines on evaluation policy:

- They need a policy which addresses both the institutional structures for managing evaluations and the openness of the evaluation process.

- They need a critical mass of evaluation staff with sufficient expertise in their field to ensure credibility of the evaluation process. Evaluations have an important role to play during a project and should not only be conducted after project completion.

- Collaboration between donors is essential in order to learn from each other and avoid duplication. Joint donor evaluations are particularly recommended.

- Aid agencies should elaborate guidelines and/or standards for the evaluation process. These should give guidance and define the minimum requirements for the conduct of evaluations and for reporting.

- Each evaluation must be planned and terms of reference drawn in order to: define the purpose and scope of the evaluation; describe the methods to be used; identify standards against which performance is to be assessed; and determine the resources and time required to complete the evaluation.

- Evaluation reporting should be clear and as free as possible of technical language. Dissemination and feedback must form a continuous and dynamic part of the evaluation process.

Source: OECD-DAC, 1991, 'Principles for Evaluation of Development Assistance', OECD-DAC, Paris
Author: Organisation for Economic Cooperation and Development - Development Assistance Committee (OECD-DAC), www.oecd.org/dac/
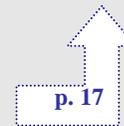⬤ Full text available online: http://www.oecd.org/dataoecd/21/41/35343400.pdf

---

4e. Standards for Evaluation in the UN System

Author: United Nations Evaluation Group (UNEG)
Date: 2005
Size: 23 pages

An effective evaluation process is an integral part of any project. But what are the key elements of a successful and sustainable evaluation approach? This document produced by the United Nations Evaluation Group (UNEG) offers solid guidelines for evaluation planning, design, implementation and reporting. Fundamental requirements include: institution-wide support, clearly-defined and transparent responsibilities, appropriately qualified staff, and a constant commitment to the harmonisation and updating of methods used.

For evaluation findings to be reliable and meaningful, they must be grounded within an established institutional framework incorporating adequate resources, standardised procedures, and thorough reporting. Conducting the evaluation itself is only one part of an ongoing process requiring stakeholder backing and the highest professional and ethical standards from all involved.

The evaluation should include an appropriate institutional framework. There should be minimum standards for the management of the evaluation function as well as a minimum set of competencies and ethics among those involved in the evaluation:

- UN organisations should have an adequate institutional framework and develop an evaluation policy to be regularly updated. They should submit evaluation plans for review and ensure appropriate evaluation follow-up mechanisms.
- The Head of evaluation should ensure that the evaluation function is fully operational, that evaluation work is conducted according to the highest professional standards, evaluation guidelines are prepared and that the evaluation function is dynamic.
- Persons engaged in designing, conducting and managing evaluation activities should possess core evaluation competencies. Evaluators should have a relevant educational background, qualification and training in evaluation, relevant professional work experience, technical knowledge of the methodology and

managerial skills. They should be culturally sensitive, respectful and should protect the anonymity and confidentiality of informants.

The design, process and implementation of the evaluation should follow the set of norms outlined below, as should the final evaluation report:

- The evaluation should provide relevant, timely, valid and reliable information. The subject, terms of reference, purpose and context of the evaluation should be clearly stated and evaluation objectives should be realistic and achievable. The evaluation design and methodologies should be rigorous. An evaluation should assess cost effectiveness and evaluation design should consider whether a human rights-based approach has been incorporated.

- The relationship between the evaluator and the commissioner(s) of an evaluation must be characterised by mutual respect and trust. Stakeholders should be consulted at all stages of the evaluation. A peer review group may be particularly useful.

- Evaluations should be conducted by well-qualified, gender-balanced and geographically diverse evaluation teams. They should be conducted in a professional and ethical manner.

- The final evaluation report should be logically structured, evidence-based, relevant, accessible and comprehensible. Evaluation requires an explicit response by the governing authorities and management addressed by its recommendations.

- The evaluation report should include an Executive Summary, complete and relevant annexes, and a clear description of the subject, context, purpose, objectives, scope, stakeholder participation, methodology and evaluation criteria. The contributions of all stakeholders should be clearly described. The report should indicate the extent to which gender issues, human rights considerations and ethical safeguards were incorporated.

- Inputs, outputs, and outcomes should be measured. Analysis should include discussion of the relative contributions of stakeholders. Constraining and enabling factors should be identified. Recommendations and conclusions need to be evidence-based, relevant and realistic, and identify important problems or issues. Lessons should have relevance beyond the immediate subject.

Source: United Nations Evaluation Group, 2005, 'Standards for Evaluation in the UN System', UNEG, New York
Author: United Nations Evaluation Group,
http://www.uneval.org/index.cfm?fuseaction=UNEG
⬤ Full text available online:
http://www.uneval.org/indexAction.cfm?module=Library&action=GetFile&DocumentAttachmentID=1496

Evaluation has a crucial role to play in today's results-based culture and in the context of the Millennium Development Goals (MDGs). How then, can the quality of evaluation be improved? This working paper from the Institute of Development Studies (IDS) argues that there has been inadequate investment in methodology, often resulting in low quality evaluation outputs. It discusses techniques in three areas of contemporary relevance: measuring agency performance; evaluation methods at the project level; and sustainability analysis.

Evaluation studies must be able to credibly establish the beneficial impact on the poor of official interventions. They must be able to draw out relevant and applicable lessons for policymakers. Evaluation in the three key areas can be improved through monitoring systems that satisfy the Triple-A criteria of aggregation, attribution and alignment, through the use of randomisation or quasi-experimental methods and by embracing the tools already available for sustainability analysis.

Many evaluation studies are data-rich but technique-poor. Problems in evaluation production result from the misuse and under-use of both data and theory:

- There is a lack of explicit attention to the techniques of meta-analysis in studies that involve the aggregation of agency ratings.

- Mean-based quantitative statistics can give a misleading summary. The best practice approach sometimes leads to the misrepresentation of data. Studies may focus on desirable processes and impacts but ignore differences in the quality of best practices or in the conditions that explain disparities in performance.

- There is often weak analysis of qualitative data, including the prevalence of data mining, sometimes formalised in the best-practice approach.

- Common approaches to the problem of attribution involve before versus after comparisons, comparisons with a control group or a combination of the two. However, such approaches have only limited applicability.

- Established methods of tackling sustainability in project appraisal suffer from being too technically sophisticated, distracting from key assumptions and from being most suitable only where variables are clearly defined.

There is thus a need to pay more attention to theory and technique, focussing on the following areas:

- Formal application of meta-analysis in studies that aggregate performance (agency-wide performance, and country and sector studies). This should be measured against the triple-A requirements of attribution, aggregation and alignment.

71

- The use of techniques to ensure that qualitative data are summarised in a way that reveals, rather than distorts, the patterns in the data.

- Paying greater attention to establishing the control in evaluation design, either through randomisation or through propensity score matching. Both techniques imply taking a prospective approach.

- Analysis of impact which is firmly embedded in a theory-based approach and which maps the causal chain from inputs to impacts.

- Seeking ways to establish impact that 'open the black box' and provide lessons about what works and what doesn't.

- Application of risk analysis to discussions of sustainability using theory-based evaluation (TBE), which seeks to uncover the key assumptions that underlie project design.

Source: White, H., 2005, 'Challenges in Evaluating Development Effectiveness', IDS Working Paper, no. 242, Institute of Development Studies, Brighton
Author: Howard White, Institute of Development Studies , http://www.ids.ac.uk/ids
Full text available online: http://www.ids.ac.uk/ids/bookshop/wp/wp242.pdf

---

4g. Strengthening the Evaluation Design and the Validity of the Conclusions

Author: M Bamberger et al
Date: 2006
Size: 24 pages

p. 18

How can threats to the validity of evaluations be identified and addressed? This chapter from Realworld Evaluation: Working Under Budget, Time, Data and Political Constraints outlines some of the most common threats to the validity of both quantitative (QUANT) and qualitative (QUAL) evaluation designs. It offers recommendations on how and when corrective measures can be taken to protect validity.

The concept of validity is closely related to that of accuracy: actual conditions must be represented in the evaluation data. In QUANT methodology, the accuracy of the data is referred to as *internal validity* or *reliability*, and in QUAL methodology, as *descriptive validity* or *credibility*. The validity of evaluation findings based on data are referred to as *interpretive* or *evaluative validity* (QUAL) and their applicability beyond the site context as *generalisability* (QUAL) and *external validity* (QUANT). The validity of an evaluation is affected by: (a) the appropriateness of the evaluation focus, approach and methods; (b) the availability of data; (c) how well the data support valid findings; and (d) the adequacy of the evaluation team to collect, analyse, and interpret data.

Evaluation designs can be assessed for potential threats to the validity of conclusions. Steps can then be taken to strengthen the likelihood of adequate and appropriate data collection and of valid evaluation findings. The Integrated Checklist for assessing evaluation validity, which includes more specific information, may be helpful.

72

To assess and strengthen QUAL evaluation designs:

- consider the comprehensiveness of data sources
- consider the cultural competence of data collectors
- consider the adequacy of ongoing and overall data analysis techniques and team capacity.

To assess and strengthen QUANT evaluation designs:

- consider whether random sample selection is appropriate and, if so, whether there is sufficient sample size or any potential sampling bias
- consider whether key indicators have been appropriately identified and whether measures or estimates of them are likely to be accurate
- consider whether statistical procedures have been appropriately selected and whether there is sufficient expertise for their use.

To assess and strengthen all evaluation designs:

- consider and use, as available, triangulation, validation, meta-evaluation and peer review
- consider the likelihood that a thoughtful combination of QUAL and QUANT approaches in a mixed-method design would improve the comprehensiveness of data and validity of findings
- consider the attitudes of policymakers and how they may affect data access and utilisation.

Source: Bamberger, M., Rugh, J., and Mabry, L., 2006, 'Strengthening the Evaluation Design and the Validity of the Conclusions' Chapter 7 in Realworld Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage Publications, California
Author: Michael Bamberger, Jim Rugh and Linda Mabry
Full text available from BLDS via document delivery

4h. Challenges in Evaluating Democracy Assistance: Perspectives from the Donor Side

Author: A T Green and R D Kohl
Date: 2007
Size: 15 pages

Why is there a lack of credible research into the impact of democracy assistance? What are the obstacles to conducting such research? This article from the journal Democratization shares insights from a donor-sponsored workshop on the challenges facing the evaluation of democracy and governance (DG) programming and assistance. It argues that the lack of credible research is partly due to a fundamental difference in orientation between the retrospective approach of academics and the prospective approach of donor agencies.

Confronting the obstacles to quality research requires understanding the political, logistical and methodological context of DG evaluation and attempting to address major issues in design and conduct. Future research needs to address a whole nexus of political issues around the appropriate role of various developing country versus donor country actors as the audience, owners, implementers, and participants in DG evaluation.

Donors need quality research on evaluating democracy assistance in order to demonstrate and assess the effectiveness of DG programming and guide future assistance. However, finding funding for evaluations in a time of shrinking budgets and increased emphasis on outputs is a major outstanding issue. There are a number of additional obstacles:

- Potentially conflicting orientations, purposes and audiences with regard to research on the effect of DG assistance and lack of resources, capacity and expertise.
- Potential bias from local experts who lack policy perspective or who focus on intended and/or more easily achieved impacts.
- Difficulties in demonstrating causality. These include deciding on the causal effect to be studied and identifying when DG assistance can be expected to have an effect.
- Making accurate and justifiable claims of attribution. Identifying programme impact in the context of multiple programmes, donors and national and global political and economic conditions presents major difficulties.
- Gaps in the coverage of data on DG, particularly for the types of countries receiving DG aid. Available baseline data is often inadequate or superficial, while the use of different indicators and classifications hinders multi-programme or multi-country research.
- Differences in programme design, content and implementation across countries, time and agencies. This makes defining a 'case' for comparative research difficult. A final methodological challenge is avoiding bias and distortion through country selection.

Donors must explicitly address these problems when designing research on DG assistance, as is the case with the following examples:

- Donor agencies from the UK, Sweden and Germany are developing alternative assessment frameworks linking project outputs to outcomes and addressing issues of causality and attribution.
- German and US agencies are developing more sophisticated quantitative analyses with improved data-sets on inputs and outcomes.
- The Netherlands Institute for International Relations (Clingendael) and the US Agency for International Development (USAID) are developing more ambitious projects using multi-programme, multi-country research frameworks.

Source: Green, A., and Kohl, R., 2007, 'Challenges in Evaluating Democracy Assistance: Perspectives from the Donor Side', Democratization, vol 14, no. 1, pp. 151-

74

165
Author: Andrew Green and Richard Kohl
⬤ Full text available online:
http://www.informaworld.com/smpp/content~content=a770216729~db=all~order=page

---

4i. Guidance for Managing Joint Evaluations

Author: DAC Evaluation Resource Centre
Date: 2006
Size: 32 pages

Joint evaluations have become central to development practice in recent years. Collective assessment of agencies' combined work minimises transaction costs for developing country partners and addresses the large aid-giving role of joined-up modalities such as basket funds and joint assistance strategies. This booklet produced by the Development Assistance Committee of the Organisation for Economic Cooperation and Development (OECD-DAC) provides practical guidance for making joint evaluations efficient, educational and collaborative.

Joint evaluations are development evaluations in which different donors and/or partners participate. The degree of collaboration can vary, falling within the following three categories: classic joint evaluations open to all stakeholders; qualified joint evaluations conducted by qualified stakeholders; and hybrid joint evaluations, involving some combination of the first two approaches.

There are five principal reasons for undertaking joint evaluations:

- Mutual capacity development occurs as agencies learn from each other and share evaluation techniques.

- Harmonisation and reduced transaction costs follow from streamlining multiple evaluations into one. At the same time, agreement can converge on evaluation messages, leading to stronger consensus.

- Participation of developing country institutions and alignment of the evaluations with national needs are generally both enhanced.

- The objectivity, transparency, and independence of evaluations are generally increased. Multiple agency participation makes follow-up on recommendations more likely.

- Broader scope for an evaluation becomes possible, facilitating a perspective on agency impacts beyond the results of one individual agency.

Despite these benefits, there are a number of potential challenges facing managers of joint evaluations. Of particular importance are the difficulty of evaluating the kinds of issues that lend themselves to joint evaluations and the complexity, cost and duration of coordination processes. Some key steps for planning and delivering joint evaluations are discussed below:

- The process of identifying evaluation partners must begin early, with research into which other agencies may be planning evaluations or may have a stake in the evaluation. Developing country partners are often ignored by donor agencies who have not prioritised evaluation capacity development or pushed for country partner budget lines for evaluations.

- Agreeing the management structure for a joint evaluation is critical. One common structure is a two-tier management system, consisting of a broad-membership steering committee and a smaller management group. Large joint evaluations can benefit from more flexible and decentralised approaches.

- Implementing and reporting on the evaluation means establishing ground rules, commonly-held terms of reference, the membership budget and the timeline of the evaluation team. This is followed by collection, analysis, and publishing of the findings.

- Dissemination and follow-up on the evaluation must be a consideration from the outset. All modes of effective communication should be explored, from traditional and internet-based publication to conferences and workshops. Follow-up strategies often include action plans and monitoring of evaluation responses, such as the Joint Evaluation and Follow-up Monitoring and Facilitation Network (JEFF) in Rwanda.

Source: OECD-DAC, 2006, 'Guidance for Managing Joint Evaluations', DAC Evaluation Series, OECD, Paris
Author: Organisation for Economic Cooperation and Development - Development Assistance Committee (OECD-DAC), www.oecd.org/dac/
Full text available online: http://www.oecd.org/dataoecd/29/28/37512030.pdf

---

4j. Sourcebook for Evaluating Global and Regional Partnership Programs: Indicative Principles and Standards

Author: Independent Evaluations Group (IEG), World Bank
Date: 2007
Size: 148 pages

Global and Regional Partnership Programmes (GRPPs) are an increasingly important modality for channelling and delivering development assistance. This World Bank Independent Evaluation Group (IEG) Sourcebook, prepared under the auspices of the OECD/DAC Network on Development Evaluation, is designed to address the growing need for consensus principles and standards for evaluating GRPPs. It comprehensively presents, synthesises, applies and elaborates on existing principles and standards, aiming to improve evaluation independence and quality. As a result, GRPPs should become more relevant and effective.

GRPPs are programmatic partnerships in which partners contribute and pool resources and establish a new organisation with a governance structure and management unit. The

76

programmes are global, regional or multi-country in scope. Most GRPPs are specific to a certain sector or theme, and almost all advocate greater attention to specific issues or approaches to development within that sector. Some programmes are primarily policy or knowledge networks, facilitating communication, advocating policy change and generating/disseminating knowledge and good practice. Others also provide technical assistance supporting national policy, institutional reform, capacity strengthening, and private/public investment. The largest programmes provide investment resources.

Notwithstanding their diversity, GRPPs have many shared features:

- They bring together multiple donors, partners and other stakeholders, whose interests may not coincide. There is joint decision making and accountability at the governance level.

- Housed in international organisations, GRPP results are the joint product of global/regional/country-level activities financed by other development agents.

- Governance and management are multi-layered and decision making is complex. Continuity may be uncertain because the members of the governing body may change due to political circumstances.

- The programme usually evolves over time, based on the availability of financing. There is not usually a fixed end-point. Decisions on which activities to support are made through a programmatic process, rather than being fixed in advance.

- GRPPs are typically externally financed with little capacity to earn income from their own resources. Financing depends on individual donors' funding decisions. GRPPs take several years to set up, and sunk costs are initially relatively high.

These shared features of GRPPs mean they require special treatment in evaluation:

- Identifying stakeholders and assessing their continued relevance is essential to determine appropriate levels of participation/consultation and possible impact on mobilisation of resources. The legitimacy and effectiveness of governance/management arrangements must be assessed.

- Evaluations must be approved by the governing body, ensuring independence and impartiality. A monitoring and evaluation (M&E) framework should be established early on. Assessing effectiveness requires consideration of the programme's inputs, outputs and outcomes at all levels. Attribution is often particularly difficult to discern in GRPPs.

- Evaluators must ascertain changes in membership of the governing body. Feedback must include all stakeholders.

- The purposes, objectives, scope and design of an evaluation must take into account the maturity of the programme. A mature programme may require exit strategies or alternative organisational/financing arrangements. The criteria and processes for resource allocation and choice of activities to support must be assessed.

- Assessing the sources/uses of funds and the relationship of the resource mobilisation strategy to the programme's scale, effectiveness and efficiency is important. Cost/benefit analysis should factor in start-up costs. Administrative costs should be assessed relative to activity costs and against external benchmarks where available.

- Improving GRPP M&E will also require much collaboration and consultation within the international development community. To that end, the wide dissemination of these principles and standards, their monitoring, and sharing experience and good practice are all encouraged amongst commissioners and providers of GRPP evaluation.

Source: IEG World Bank and OECD-DAC Network on Development Evaluation, 2007, 'Sourcebook for Evaluating Global and Regional Partnership Programs: Indicative Principles and Standards', Independent Evaluation Group – World Bank, Washington, D.C.
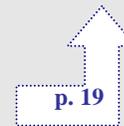**Author:** Christopher D Gerrard, www.worldbank.org
 Full text available online: http://www.oecd.org/dataoecd/38/54/37981082.pdf

---

4k. Reconciling Different Priorities and Perspectives: Addressing Political Influences

Author: M Bamberger et al
Date: 2006
Size: 19 pages

p. 19

No evaluation can ever be value free and completely objective. Decisions as to what to study, which methods to use, and whose criteria define programme success all involve human judgement. This chapter from RealWorld Evaluation: Working Under Budget, Time, Data and Political Constraints, discusses how political factors affect evaluation. It provides a detailed analysis of possible pressures and constraints in evaluation design, implementation, dissemination and use.

Development programmes are part of a political agenda. Evaluation can be extremely challenging, as it may confirm or confront the programme concerned, its underlying values, supporters and opponents. Political manoeuvring is neither bad nor avoidable – it signals the importance of evaluation – but evaluators should be alert to political and ethical issues from the outset. Societal politics have, at times, proved critical in the history of evaluation. Early evaluations were mainly financial and quantitative. The past decade has seen an increasing emphasis on participatory evaluation and the involvement of a broad range of stakeholders.

Some features of the divergent perspectives found in evaluation are outlined below:

- It is not easy to identify the multiplicity of stakeholders' values, nor to prioritise their interests appropriately. Some stakeholders wield greater influence and can

use their position to affect evaluation focus, criteria, methodology or data access. Stakeholder dissension, rather than consensus, may be likely.

- Individual evaluators hold diverse professional values regarding the evaluator's role. This can sometimes obstruct the development of external accreditation or licensing.

- Evaluators will inevitably be influenced by their own perspective and values when determining a programme's merit or quality, even when discerning what constitutes 'data'.

- Codes of professional conduct can be helpful in clarifying potential responses to political situations and in defending those responses to clients and stakeholders. However, in complex scenarios, these codes do not provide clear priorities or choices, but rely on individual judgments.

Programme evaluations are often commissioned with political motives. At all stages of an evaluation, the evaluator may encounter hidden agendas, obstruction, conflicting priorities, and his or her own shifting role – as guide, publicist, friend and critic. Evaluators may adopt the following strategies to address political constraint:

- During the defenition of the terms of reference for the evaluation: The client's choice of evaluator can signal political intention. Negotiation before the evaluation starts may establish helpful ground rules.

- During the evaluation design: Understanding the political environment may help an evaluator identify ways to address the pressure exerted on the evaluation. Stakeholder analysis should be conducted to identify priorities, concerns and constraints. Participatory planning and consultation may help create shared investment in the evaluation.

- During the evaluation implementation: Good communication may facilitate the evaulator's access to sources of information which the client or other stakeholders were initially reluctant to make available. It can help to identify the data required, provide feedback to allay suspicion, and demonstrate the value of the evaluation.

- In the presentation and use of evaluation findings: Evaluators sometimes encounter clients who do not intend to publish or use findings accurately. An effective technique to guard against this is to ensure findings are of direct practical utility to the different stakeholders.

Source: Bamberger, M., Rugh, J., and Mabry, L., 2006, 'Reconciling Different Priorities and Perspectives: Addressing Political Influences', Chapter 6 in Realworld Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage Publications, California.
Author: Michael Bamberger, Jim Rugh and Linda Mabry
⬤ Full text available from BLDS via document delivery

4l. Influential Evaluations: Evaluations that Improved Performance and Impacts of Development Programs

Author: World Bank Operations Evaluation Department
Date: 2004
Size: 24 pages

Evaluations can be a cost-effective way of improving the performance and impact of development activities. However, they must be conducted at the right time, focus on key issues and present results in an accessible format. This report from the World Bank Operations Evaluation Department presents eight examples of evaluations that have had an important impact, and summarises lessons learned.

The money spent on evaluations is only justified if they are relevant to policy makers and managers and help precipitate change. For this, they must be timely, targeted and easy to use. Evaluations that fail to meet these criteria may not produce useful results, even if their methodology is sound. By outlining eight case studies, the report sheds light on the design of useful evaluations and how their utilisation and cost-effectiveness can be measured. The evaluations featured assess interventions to: improve the efficiency of India's employment assurance; use citizen report cards to hold the state to account in Bangalore, India; improve water and sanitation in Flores, Indonesia; broaden the policy framework for assessing dam viability (World Bank); abolish ration shops in Pakistan; improve the delivery of primary education in Uganda; enhance the performance of a major environmental project in Bulgaria; and re-assess China's national forest policy.

The report identifies a number of general factors that make an evaluation more likely to enhance the performance and impact of development interventions:

- A conducive policy environment. Findings are more often used when they address key concerns and decision-makers accept the consequences of implementation.
- The timing. The evaluation should be launched when there is a clear need for information, and findings delivered in time to affect decisions.
- The role. Evaluations are usually one of many information sources for policy makers, and should be adapted to the context in which they are used.
- Relationships and communication of findings. A good relationship should be developed with stakeholders and they should be informed of progress so that they are not surprised by the results.
- Who conducts the evaluation. This could be the evaluation unit of the managing/funding agency (which has access to actors and data), an outside organisation (which may be more objective) or a joint team.

There is no single best methodology for conducting evaluations. The right approach will depend on the context, issues to be addressed, and the resources available. Most combine different methods to increase the reliability of findings and to build a broader framework for their interpretation. Other lessons are that:

- the value of an evaluation should be assessed in terms of its cost-effectiveness. An 'expensive' evaluation is justified if it produces cost reductions or produces benefits that significantly outweigh its own cost.
- where benefits cannot be judged in monetary terms, decision-makers must judge whether the investment in the evaluation will produce adequate results.
- assessments of evaluations are usually constrained by time, budget and political constraints. This means they are often less rigorous than recommended.
- utilisation can be assessed by reviewing reports, discussions with the evaluation team and asking the opinion of local experts.
- assessing how much an evaluation has contributed to outcomes (attribution analysis) is difficult, but all available tools should be used.
- the assumptions and estimation methods of cost-effectiveness analysis should be clearly stated so that their validity can be appraised.

Source: World Bank Operations Evaluation Department, 2004, 'Influential Evaluations: Evaluations that Improved Performance and Impacts of Development Programs', World Bank, Washington, D.C.
Author: Independent Evaluations Group, World Bank, http://www.worldbank.org/oed
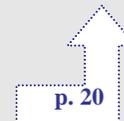
🔴 Full text available online:
http://lnweb18.worldbank.org/oed/oeddoclib.nsf/24cc3bb1f94ae11c85256808006a0046/67433ec6c181c22385256e7f0073ba1c/$FILE/influential_evaluations_ecd.pdf

---

4m. From Policy Lessons to Policy Actions: Motivation to Take Evaluation Seriously

Author: G Gordillo and K Andersson
Date: 2004
Size: 16 pages

p. 20

Whilst recent political reforms have sometimes led to modifications in countries' national policies, the link between policy evaluation and policy actions is often weak. So why do so many governments take policy evaluation so lightly? This article from the Food and Agriculture Organisation (FAO) and the Center for the Study of Institutions, Population, and Environmental Change (CIPEC) at Indiana University analyses the institutional aspects of creating effective systems for monitoring and evaluations (M&E) in government-led rural development efforts in Bolivia and Brazil.

The difficulty with transforming M&E information from a restricted 'club good' into a public good relates to motivation problems and power and information asymmetries. The mere existence of robust policy lessons is not sufficient to generate policy actions. Political pressure and potential financial gains could potentially counteract unproductive institutional arrangements. However, translating policy lessons to policy actions requires decision-makers to change their attitudes toward evaluations. This is unlikely to occur

81

without successively strengthening the institutions for equitable political representation and downward accountability.

The Institutional Analysis and Development (IAD) framework is used to examine the incentive structures of the actors involved in rural development efforts in Bolivia and Brazil. In both cases, the establishment of mechanisms for downward accountability was crucial in motivating politicians to take the results of the studies seriously:

- Despite the introduction of Bolivia's 1994 Law of Popular Participation, local community groups saw little improvement in access to forest management rights. Most of those groups that *were* granted rights had exerted considerable political pressure at high political levels.
- The general failure of community forestry concessions in Bolivia resulted from a lack of undisputed public forest land that could be considered for community concessions, a cumbersome and bureaucratic process and low motivation among policy makers to meet the demands of rural populations.
- Nevertheless, the successful cases illustrate the power of M&E systems when local user groups learn how to use its results to hold politicians accountable for their policy actions, or lack thereof.
- In Brazil, the government contracted rural development specialists to address failures in its programme of family agriculture and subsequently opened up a period of extensive public consultation.
- Several of the resulting proposals were adopted. This was motivated by a dramatic, unequivocal and timely message, a desire to remedy failed attempts to address land concentration, broad consultation and involvement from FAO.

The cases from Bolivia and Brazil point to the importance of considering the motivation of politicians to respond to evaluations:

- Producing good quality evaluation studies alone is not sufficient to motivate political action. This requires institutional incentives that are compatible with action and change. In addition, beneficiary populations and important interest groups need to play a role in evaluation design and implementation.
- Evaluation should be promoted as an instrument for downward accountability through development projects supported by international cooperation agreements. Both NGOs and intergovernmental organisations have important roles to play in facilitating the participation of marginalised groups in the policy process. A concerted effort among these actors has the potential to connect the feedback loops from evaluation to action.
- Evaluations need to be used as a tool for learning. This can lead to more reliable information from evaluations and can increase the possibilities of learning lessons about the effects of earlier policies.

Source: Gordillo, G., and Andersson, K., 2004, 'From Policy Lessons to Policy Actions: Motivation to Take Evaluation Seriously', Public Administration and Development, vol 24, no. 4, pp. 304-320

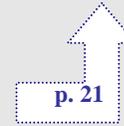**Author:** Gustavo Gordillo and Krister Andersson

---

4n. Building Country Capacity for Monitoring and Evaluation in the Public Sector: Selected Lessons of International Experience

Author: World Bank
Date: 2005
Size: 22 pages

The Evalution Capacity Development (ECD) unit of World Bank's Operations Evaluation Department is designed to help countries strengthen their monitoring and evaluation (M&E) capacity. The unit targets 'high-intensity' support to Uganda and Egypt and various other types of support to an additional 32 countries. This paper from the World Bank collates some of the main lessons learned from ECD activities and outlines the major issues which need to be addressed.

Experience shows that almost any regime can produce short-lived results, but only governments with a sound basis of legitimacy can produce sustainable ones. Because, in the public sector, this can only be verified after many years, attention must be paid to the governance context within which M&E capacity is to be built. Strong governance, accountability and transparency must all be considered when evaluating the likelihood of the long-term sustainability of public management policies and programmes.

In developing countries, M&E should be introduced at an early stage in the reform process. A good understanding of local realities, particularly institutional and administrative capacity, is required to accurately define an appropriate scope and timeframe for ECD assistance. The following issues must be addressed early in the implementation of ECD activities:

- Evaluation may be approached through assessment of met objectives or of achieved results. Objective assessment is required for the long-gestating impact of M&E capacity improvements, complemented by mid-term assessments and close collaboration between M&E capacity efforts and donors.

- Government in-house or independent M&E capacity? Whilst both have their strengths and weaknesses, experience shows that thorough evaluations require substantial resources which are lacking in most developing countries' governments. Creating M&E capacity requires evaluation capacity fostered outside of government and direct help to create strong in-house capacity to design, guide and monitor the external evaluator.

- The interaction between the executive and the legislature is important for future M&E capacity building, both in terms of the legislative capacity to impact intelligently on budget formulation and the assessment of budget execution.

- Performance measurement is a means and not an end. Performance indicators must be simple and clear and chosen with participation from front-line staff and service-users. Their effectiveness should be assessed regularly and adjusted. Expansion must incorporate feedback.
- Most ECD work to date has focused on macro-level issues, but support for M&E depends on visible results on the ground, requiring improvement at sector ministry level, and even at the level of specific public services.
- The preponderance of evidence demonstrates the importance for developing countries to focus on the step-by-step problem analysis of M&E capacity. However, there must also be a larger vision of what ECD 'success' looks like.

Building effective M&E capacity is neither quick nor easy:

- Even in highly advanced countries, the evaluation system suffers from some of the same pitfalls evident in developing countries: weak piloting, overstated achievements, ambiguous results and timing discontinuities.
- The current systems of M&E and programme reviews emerged from almost 30 years of experimentation, learning, and required sustained efforts at strengthening institutional, organisational and human capacity.
- Steady and sustained support from international donors and intelligent and realistic sequencing of ECD assistance is essential to the process.

Source: Schiavo-Campo, S., 2005, 'Building Country Capacity for Monitoring and Evaluation in the Public Sector: Selected Lessons of International Experience', World Bank Evaluation Capacity Development Working Paper, no. 13, World Bank, Washington, D.C.
Author: Salvatore Schiavo-Campo, World Bank , www.worldbank.org
 Full text available online:
http://lnweb18.worldbank.org/oed/oeddoclib.nsf/24cc3bb1f94ae11c85256808006a0046/d0833fcf23ecceeb8525702f0059bd7e/$FILE/ecd_wp_13.pdf

---

5a. Reversing the Paradigm: Quantification and Participatory Methods

Author: L Mayoux and R Chambers
Date: 2003
Size: 25 pages

What role should participatory methods play in assessing the impact of development activity? A common assumption is that rigorous quantitative data can only be generated by questionnaire surveys or scientific measurement. Another is that participatory methods can only generate qualitative insights. This paper from the Enterprise Development Impact Assessment Information Service (EDIAIS) discusses experiences and innovations which show these assumptions to be false. It argues that participatory approaches can generate accurate qualitative and quantitative information and should form the basis for monitoring, evaluation and impact assessment.

84

Participatory methods for generating numbers have proved more cost-effective than conventional approaches. They can also be used as a basis for targeting and focusing more expensive quantitative and qualitative investigations if these are still needed. Although empowerment of participants cannot be assumed, participatory methods have substantial potential to contribute to civil society development and to downward accountability in development processes.

Yet, wider use of participatory methods is inhibited by institutional inertia and professional conservatism. More specific challenges include potentially biased discussions and difficulties with systematic sampling and analysing information. Furthermore, outsiders often have difficulty understanding diagrams. Overcoming these weaknesses will enable the poor to have more voice in priorities, policies and practices.

Despite these challenges, participatory methods give access to information that would be difficult to obtain through conventional quantitative approaches:

- Conventional quantitative methods generally fall short of impact assessment requirements. Failure to cover non-economic facets of poverty and sample sizes too small to allow reliable statistical inferences make the quantification of impacts less rigorous than is often claimed.

- Participatory methods – using focus group discussions and diagram tools – have the advantage of cost-effectiveness in rapidly gathering information from many participants. Information is likely to be more reliable due to immediate verification from other participants.

- Numerous methods can be used in group meetings to generate numerical data, even for issues requiring anonymity. Methods include shows of hands, secret ballots, pocket voting and participants plotting themselves on diagrams and grouping themselves by characteristics.

- Quantitative data can come from measuring, counting, estimating, valuing, ranking and scoring. There have been many cases of information generated by group exercises being aggregated over a whole area.

Thus, participatory methods should be considered ahead of conventional quantitative survey methods for impact assessment:

- Participatory methods are particularly effective for rapidly collecting information which is common knowledge and visually cross-checked in a group process. They are also particularly effective where people have information on different parts of a whole.

- Participatory methods are as effective as quantitative methods, and often more so, where people wish to give false information or information is disputed.

- Individual qualitative interviews may be necessary to discuss personally sensitive issues.

- Questionnaire surveys could be reserved for ascertaining the generalisability of sensitive qualitative issues and circumstances where representation in participatory processes is highly skewed.

- While increasing the rigour and scale of information from participatory processes, empowerment and ownership of assessment and development processes remains a constant concern.

Source: Mayoux, L., and Chambers, R., 2003, 'Reversing the Paradigm: Quantification and Participatory Methods', Enterprise Impact, London
Author: Linda Mayoux and Robert Chambers, Enterprise Development Impact Assessment Information Service , http://www.enterprise-impact.org.uk/
⬤ Full text available online:
http://www.enterprise-impact.org.uk/word-files/Chambers-MayouxUpdate.doc

---

5b. Problematizing Participation: A Critical Review of Approaches to Participation in Evaluation Theory

Author: A Gregory
Date: 2000
Size: 22 pages

It is widely accepted that evaluation is a social process which implies the need for a participatory approach. But what is understood by 'participation'? This critical review from Hull University Business School argues that the blanket use of the term has masked the heterogeneity evident in its realisation in practice and highlights a lack of transparency in participatory methods in evaluation.

Four key arguments are typically advanced to support the notion of participation in decision-making, design and planning: ethics, expediency, expert knowledge and motivating force. Yet the notion of participation is ill-understood and is an important problem across a range of methodologies in evaluation. Participation should be explicitly considered rather than ignored or implicitly assumed. In this sense, the systems field may offer guidance on how best to realise a participatory approach. The problem of participation can only be approached through an understanding of power and its realisation in practices that prohibit or promote participation.

The lack of evident success in practice is a result of a general failure to make the nature and aims of participation explicit. To what extent are evaluation methodologies guilty of this?

- Rebien's 'Continuum and Criteria' are based on different degrees of participation. He advances a continuum of participation with a threshold to identify 'truly' participatory projects. Yet Rebien's criteria for distinguishing participatory methodologies are insufficiently defined and may promote practices which impact negatively on participation.
- Guba and Lincoln's 'Fourth Generation Evaluation' is analysed through Oakley's obstacles to participation. Whilst explicitly participative in nature, it includes some rather naïve assumptions which are unlikely to promote participation.

86

- Patton's 'Utilisation-Focused Evaluation', grounded in realism, advocates limited composition and size of the evaluation task force. Patton's ideas about methodological transparency and capacity building aspirations are unrealistic and his approach may result in only limited participation.

- Pawson and Tilley's 'Realistic Evaluation' seeks to explain how mechanisms operate in contexts to produce outcomes. It is based on the separation of roles and on a limited form of participation. Consequently, the knowledge produced is critically restricted.

- Taket and White's 'Working with Heterogeneity' stems from systems analysis. They advocate a form of evaluation based on 'pragmatic pluralism'. However, their approach may promote an expert driven form of evaluation. Whilst acknowledging the influence of power on participation, their recommendations for dealing with power are not well developed.

Only through an appreciation of power can the problem of participation be addressed. Participatory methodologies must be simple if they are to be transparent and easily transferable. As a starting point, Ledwith's 'Sites of Oppression' matrix is recommended as a good way of exploring how the processes of power operate:

- The matrix illustrates the potential ways in which oppression overlays and interlinks. It includes elements of oppression and the different, reinforcing and interdependent levels at which oppression operates. It facilitates a deeper analysis and offers a wider perspective.

- Such an analysis would benefit those of a pluralist orientation by clarifying who should be involved in the evaluation, what barriers exist to prevent participation, and how these might be removed.

- Those with a more realist bent would gain an appreciation of how mechanisms/structures work at different levels. This analysis would draw on the propositional, experiential and practical knowledge of the subjects of evaluation.

Source: Gregory, A., 2000, 'Problematizing Participation: A Critical Review of Approaches to Participation in Evaluation Theory', Evaluation, Vol 6, no. 2, pp. 179-199
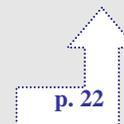**Author:** Amanda Gregory, Hull University Business School, http://www.hull.ac.uk/hubs/
Full text available online: http://evi.sagepub.com/cgi/content/abstract/6/2/179

5c. Learning from Change: Issues and Experiences in Participatory Monitoring and Evaluation

Author: M Estrella
Date: 2000
Size: 15 pages

Since the 1980s concepts of participatory monitoring and evaluation (PM&E) have entered the policy-making domain of larger donor agencies and development organisations. This introductory chapter from Learning from Change: Issues and

Experiences in Participatory Monitoring and Evaluation draws on twelve case studies to describe how different stakeholders have applied PM&E approaches across a range of purposes and contexts. It outlines some of the key concepts and differences between participatory and conventional approaches to M&E and highlights some emerging issues.

The conventional 'top-down' and 'objective' approach to M&E which excludes many stakeholders has led to dissatisfaction in the international development community. Growing interest in PM&E is a reflection of several emerging trends: 'performance based accountability'; increasing fund scarcity and pressure to demonstrate success; and the shift towards decentralised and devolved governance requiring new forms of oversight. Meanwhile, the increasing capacity and experience of non-governmental organisations (NGOs) and community-based organisations (CBOs) as decision-makers and implementers in the development process provide additional impetus. Across a varied range of PM&E methods, four common features of good practice are identifiable: participation, learning, negotiation and flexibility.

There are four key stages in establishing the PM&E process – planning the framework and determining objectives and indicators; gathering data; analysing and using data by action-taking; and documenting, reporting and sharing information. The critical feature of a PM&E approach is its emphasis on who measures change and who benefits from learning about these changes. Stakeholders directly involved in a programme take part in selecting indicators, collecting information and evaluating findings. Negotiating and resolving different stakeholders' needs in order to make collaborative decisions remains a critical question in building the process. PM&E provides information to meet different stakeholder needs and objectives:

- Project planning and implementation - Have project objectives been met? How have resources been used?

- Organisational strengthening and institutional learning - This involves enabling NGOs, CBOs and people's organisations to keep track of progress, build on areas of successful work and develop sustainability. Enabling local stakeholders to measure institutional performance fosters social accountability and responsiveness.

- Informing policy - Where local communities are empowered to communicate local needs, these may be compared against local government development priorities.

By encouraging stakeholder participation beyond data-gathering, PM&E is about promoting self-reliance in decision-making and problem-solving, strengthening people's capacities to take action and promote change. Whilst some common guidelines are emerging which help define how PM&E is established and implemented, several issues require further exploration:

- Clarifying concepts of participation – There remains great ambiguity in defining who stakeholders are and to what extent they should be involved, overall, and at different stages of the process.

88

- Identifying appropriate methodologies – Procedures for indicator development are not always clear when incorporating differing stakeholders' needs and priorities. The more context-specific information obtained in PM&E approaches is perceived as 'subjective'. The question remains whether it is therefore less rigorous.

- Developing capacity building – There is recognition that PM&E requires considerable time, financial investment and human resources, but little documentation to identify the necessary requirements to build and sustain PM&E over time.   What types of skills, knowledge, attitudinal and behavioural changes are required to conduct PM&E?

- Scaling up PM&E and promoting institutional learning – Can PM&E be built into the standard operating procedures of formal institutions?

Source: Estrella, M., 2000, 'Learning from Change', in Learning from Change: Issues and Experiences in Participatory Monitoring and Evaluation, ed. M. Estrella, Intermediate Technology Publications and the International Development Research Centre, London and Ottawa
Author: Marisol Estrella, International Development Research Centre (IDRC), http://www.idrc.ca
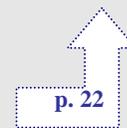🟠 Full text available online:
http://siteresources.worldbank.org/INTPCENG/1143331-1116505657479/20509240/learnfromchange.pdf

---

5d. The Power of Participatory Monitoring and Evaluation: Insights from South-West China

Author: R Vernooy
Date: 2006
Size: 12 pages

Participatory monitoring and evaluation (PM&E) is a partnership between researchers and other stakeholders to systematically assess research or development activities. Focusing on participatory field research for community-based natural-resource management (CBNRM), this article from *Development in Practice* describes the capacity-building experiences of two research teams in Yunnan and Guizhou provinces in south-west China. The ongoing democratisation and decentralisation processes in China aim to allow more space for local voice and decision-making power over NRM. So who participates and what difference does participation make?

Conventional M&E mainly serves the needs of project implementers and donors, ignoring the interests of other stakeholders, especially local people. The participatory process may increase the relevance and effectiveness of research and encourage local empowerment and social transformation. This can help strengthen people's capacity to make decisions and create an environment for change, although some tensions and unintended consequences are inevitable. In China, local politics – from village, to township, to country levels – play a key role in the process of rural change.

The PM&E process undertaken by the CBNRM teams contributed to a better understanding of how the concerns of stakeholders are represented and negotiated in a research process.

- The participants (farmers, government officials and researchers) developed a greater degree of trust, understanding and co-operation to achieve shared goals. Overall, communities became stronger.

- Farmers (particularly women) were better able to participate in the research and change process and had more ownership of the projects' processes and outcomes.

- Project management improved, with more space for reflection, responsiveness and adaptation. The projects' objectives and outcomes moved closer to local needs.

- Researchers and farmers built a solid relationship over time and focused attention on villagers' immediate livelihood needs. This enabled locals to overcome past experiences of marginalisation. However, the time required to achieve this level of participation poses a challenge for the future.

PM&E was important in identifying problems, opportunities and strategies, and building capacity, accountability and confidence in the projects. Some recommendations can be made for the structure of future PM&E programmes and their likely impact on local projects:

- Quality is improved by integrating a PM&E system into project management from the planning stages. The practices of participation and decision-making empowerment both need improvement particularly in considering gender and other social variables.

- PM&E should be institutionalised at each level of project management and all stakeholders should understand its benefits. PM&E must be built on existing community institutions and integrated into local governance structures and political processes.

- PM&E spreads the risk of failure between the project team and participants – when the project bears all the risk, local participants care less about their own role in the project.

- Where stakeholder representatives are determined through locally democratic means, PM&E can help to promote democracy, decision-making, local management and accountability.

- PM&E can raise local awareness about rights and responsibilities and good governance principles amongst local officials. PM&E may also encourage township and county-level governments to experiment with local solutions and greater autonomy in villages.

Source: Vernooy, R., Qui, S., and Jianchu, X., 2006, 'The Power of Participatory Monitoring and Evaluation: Insights from South-West China', Development in Practice, Volume 16, Number 5, pp. 400-411
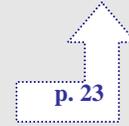**Author:** Ronnie Vernooy, International Development Research Centre (IDRC), http://www.idrc.ca

Full text available from BLDS via document delivery

5e. Participatory Monitoring and Evaluation: Documentation to Support Generic Framework

Author: ITAD and RuralNet Associates
Date: 2004
Size: 66 pages

The Participatory Monitoring and Evaluation (PM&E) consultancy for the Zambia Social Investment Fund (Zamsif) aims to establish sustainable PM&E processes at both the community and district levels. This report from ITAD and RuralNet Associates discusses the development and rationale of the PM&E framework chosen for the project. It constructs detailed diagrams to model the consultancy's work and explains current and potential uses of the framework.

The specific objectives of the consultancy are to: design a framework for developing relevant indicators and methods; develop frameworks for reflective learning and other stages in the design of PM&E systems at all levels; develop training manuals in participatory monitoring and evaluation; and conduct relevant training in PM&E for Zamsif staff.

The Zambian context features a government-initiated poverty reduction scheme taking a Community Driven Development (CDD) approach. This approach aims to devolve control of decisions and resources to local units. PM&E has been shown to enhance accountability, empowerment, and strong partnerships between the key stakeholders. Core principles of participation, negotiation, learning, flexibility, and the involvement of stakeholders were salient for the consultancy. A particular innovation of Zamsif is a district graduation programme that grants additional resources and powers to districts that achieve performance benchmarks. Key elements of the framework are outlined below:

- The genesis of the framework began with a national workshop that developed a number of areas and sub-areas deemed important by participants and identified potential indicators for the M&E process.

- The next step involved district-level and community-level consultations, as well as piloting in two districts - Chipata and Chibombo. Each of these steps strengthened the framework and contributed key recommendations.

- The framework is based on straightforwardness, simplicity, flexibility and adaptability and includes a checklist of themes to spur innovative thinking. Building on previous team experience, developing a realistic action plan, and integrating the M&E with other management tools are also key.

- A successful PM&E system requires the timely and relevant flow of information. Information sharing between levels should be multi-directional to provide districts and communities with regular and frequent information on project progress.

- Reflective learning is undertaken to identify lessons learned from practical experience and based on PM&E data. This contributes to improved planning as well as project and programme development. The process also allows individuals

to reflect on their own practices and helps build a learning environment within the organisation.

The consultancy's generic framework formed the basis for training materials, including a Trainers' Guide and course curriculum, a District Implementation Manual and a Community Facilitators' Guide. A number of recommendations are made for the remainder of the consultancy:

- Creating an institutional home for PM&E skills
- Ensuring that training materials are available at district level
- Providing refresher training and workshop support for PM&E training
- Identifying constraints experienced at the district and community levels
- Supporting direct negotiations of communities and districts
- Monitoring the timeliness, usefulness, and appropriateness of multi-level information flow
- Ensuring that the fund's documentation is made compatible with PM&E documentation
- Bolstering central support for the consultancy.

Source: ITAD and RuralNet Associates, 2004, 'Participatory Monitoring and Evaluation: Documentation to Support Generic Framework', ITAD, London
Author: ITAD, http://www.itad.com/
⬤ Full text available online:
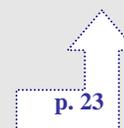http://www.itad.com/neweb/pubs/PM&E%20Framework%20AUG04.pdf

---

5f. From PRA to PLA and Pluralism: Practice and Theory

Author: R Chambers
Date: 2007
Size: 41 pages

How have the theory and practice of participatory methodologies in development activities changed since the mid 1970s? What variants and applications of these methodologies have emerged? This paper from the Institute of Development Studies (IDS) traces the spread of Participatory Rural Appraisal (PRA) and Participatory Learning and Action (PLA), and identifies strengths and weaknesses in the ways they have been implemented in development contexts. Whilst inflexible applications of PRA and PLA may produce disappointing results, when executed with spontaneity and creativity, these approaches can be a source of community empowerment.

PRA and the more inclusive PLA are families of participatory methodologies which have evolved as behaviours and attitudes, methods, and practices of sharing. More recently, PRA has also come to mean Participatory Reflection and Action, as a result of shifts in its practical application. The term Participatory Learning and Action (PLA) is broader than

92

PRA and includes other similar or related approaches and methods. Because of the continuities and overlaps, this methodological cluster is often referred to as PRA/PLA.

In the evolution of PRA/PLA, there was much intermingling and innovation. Among other sources were the approaches and methods of action science, reflection-in-action, popular education, and participatory research and participatory action research. Beyond this illustrative listing, more of a sense of what has happened can be given through eight examples of parallel and intermingling PRA/PLA:

- Farmer Participatory Research programmes have shifted from an emphasis on researcher design and control to a farmer-led approach which has substantially improved outcomes.

- Integrated Pest Management, which trains Indonesian farmers to control pests through action research, has provided a safe space for social learning and action even within a repressive social order.

- 'Reflect' applies the participatory approach to literacy and empowerment processes and has spread to more than 60 countries, working best when programmes are adapted to their social context.

- Stepping Stones aims to facilitate experiential learning about social awareness, communication and relationships through group interaction and has received near universal support from participants.

- Participatory Geographic Information Systems combine PRA/PLA techniques and spatial information technologies to empower groups traditionally excluded from geographical decision-making processes.

- The Internal Learning System incorporates pictorial diaries to enable poor, often illiterate participants to reverse normal power structures by gathering data on social change.

- Participatory Action Learning Systems provides individuals with the tools to collect and analyse the information they need to improve their lives in areas such as women's empowerment.

- Community-Led Total Sanitation offers individuals the opportunity to research their own defecation practices in order to improve local sanitation services.

From these examples, some general principles can be drawn out for applying PRA/PLA methodologies:

- Practitioners should remain humble before local people's own knowledge.

- Improvisatory approaches in the field can provide opportunities to enrich the practice of PRA/PLA methodologies.

- Simply formulated precepts can effectively embed changed attitudes and behaviours in the minds of participants.

- Visual, tangible methods such as modelling or scoring with seeds can encourage collective analysis, represent complex realities and empower marginalised individuals.

93

- Projects governed by very few rules can often prompt complex practices which embody local people's creative and instructive responses to their surroundings.
- Participatory methodologies should remain eclectic and inventive in order to encourage sharing of knowledge and avoid the trap of ownership and branding.

Source: Chambers, R., 2007, 'From PRA to PLA and Pluralism: Practice and Theory', IDS Working Paper, no. 286, Institute of Development Studies, Brighton
Author: Robert Chambers, IDS, www.ids.ac.uk
Full text available online: http://www.ids.ac.uk/ids/bookshop/wp/wp286.pdf

---

5g. Using Key Informant Monitoring in Safe Motherhood Programming in Nepal

Author: N Price and D Pokharel
Date: 2005
Size: 14 pages

p. 24

The Nepal Safer Motherhood Project (NSMP) works to improve maternal health and contribute to programme development at district and national level. This article from Development in Practice discusses the project's use of Key Informant Monitoring (KIM). KIM is an adapted version of the peer ethnographic research method. Data is collected by community-based Key Informant Researchers (KIRs) and used for monitoring and planning. KIRs have proved useful sources of information and acted as change agents by spreading safer motherhood messages.

Only 13% of births in Nepal are attended by a health professional and the unmet need for obstetric care is estimated at 95%. Using KIM, NSMP aims to understand the socio-cultural, economic and political environment in which pregnancy and childbirth are experienced and monitor progress in improving access. KIM takes as its starting point the idea that this social context is important in shaping health-seeking behaviour and maternal outcomes. Developing and implementing KIM has led to recognition among NGOs of its potential for use in other activities associated with safe-motherhood, HIV and reproductive health.

KIM is an adaptation of the peer ethnographic method, which draws on the participant-observation approach of anthropology, encouraging trust and rapport between the researcher and the researched.

- Rural Nepal is highly stratified along lines of ethnicity, gender, kinship and age. These divisions and power relations must be considered when exploring community perceptions.
- While there are no clearly defined peer groups in rural Nepal, KIM has sought to train women to interview others of similar age and social background.
- There are limits to the amount of information KIRs are able to record. Instead of an interview script, KIRs make use of conversational prompts to collect data on barriers to services, quality of care and women's decision making.

94

- After gathering data, KIRs attend debriefing workshops with NSMP female local facilitators. Data analysis also involves review and dissemination workshops led by NGOs at Village Development Committee (VDC) headquarters.

- Steps taken by VDCs in response to KIM findings are further reviewed and refined by subsequent rounds of KIM data collection. Findings are also shared with the District Reproductive Health Co-ordination Committees.

- Responses to KIM findings have included: awareness raising campaigns, training for traditional healers, better privacy in examination rooms, more emergency funds and greater involvement of women in VDCs.

Key lessons and implications relate to local ownership, partnership and methodology:

- KIM has facilitated participatory dialogue between NSMP, its key partners and the community. It has fostered community ownership of methods and data and produced credible findings.

- Through their interaction with women and their families, KIRs have acted as catalysts for dialogue and behaviour change, for example, by convincing family members to take women with obstetric complications to hospital.

- One methodological issue that had to be addressed early on was possible KIR bias. Some KIRs had become over-committed to NSMP objectives and overzealous in the pursuit of evidence of better access and service quality.

- The generic peer ethnographic method lends itself to use in urban settings, where there is a sense of anonymity. Its use in rural Nepal required adaptation, but ulitimately demonstrated how the method could be applied in different contexts.

Source: Price, N., and Pokharel, D., 2005, 'Using Key Informant Monitoring in Safe Motherhood Programming in Nepal', Development in Practice, Volume 15, Number 2
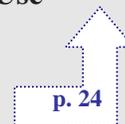**Author:** Neil Price and Deepa Pokharel, Centre for Development Studies, Swansea (CDS), http://www.swan.ac.uk/cds
Full text available from BLDS via document delivery

5h. The 'Most Significant Change' (MSC) Technique: A Guide to Its Use

Author: R Davies and J Dart
Date: 2005
Size: 104 pages

This paper from MandE News, an online monitoring and evaluation news service, outlines an innovative qualitative monitoring technique known as the 'most significant change' (MSC) approach. The MSC technique is a participatory method of collecting and analysing stories from the field which focuses on monitoring intermediate outcomes and impact. It provides a simple means of making sense of a large amount of complex information and is best suited to large-scale, open-ended projects which would otherwise be difficult to monitor easily using traditional methods.

The MSC process involves the following steps: raising interest; defining the domains of change; defining the reporting period; collecting significant change (SC) stories; selecting the most significant of the stories; feeding back the results of the selection process; verifying the stories; quantification; secondary analysis and meta-monitoring; and revising the system.

MSC monitoring is useful for identifying unexpected changes. It requires no special professional skills, encourages analysis as well as data collection and can build staff capacity. It can deliver a rich picture of what is happening and can be used to monitor and evaluate bottom-up initiatives that do not have predefined outcomes.

MSC is better suited to some programme contexts than others and has a number of advantages and drawbacks compared to other forms of M&E:

- MSC is suited to monitoring that focuses on learning rather than just accountability. The types of programmes can gain considerable value from MSC include those that are complex, large, focused on social change, participatory and highly customised.

- MSC may be less appropriate for: capturing expected change; developing good news stories; conducting retrospective evaluation; understanding the average experience of participants; producing an evaluation report for accountability purposes; or for completing a quick and cheap evaluation.

- MSC helps draw valid conclusions through thick description, systematic selection, transparency, verification, participation, and member checking.

- Some of the key enablers for MSC are: an organisational culture where it is acceptable to discuss failures; champions with good facilitation skills; a willingness to try something different; infrastructure to enable regular feedback; and commitment by senior managers.

- Problems with MSC relate to the meaning, significance and relevance of the question, the selection of SC stories, time constraints, and complaints that certain choices are ignored and feedback forgotten. Furthermore, MSC contains a number of biases as well as subjectivity in the selection process.

MSC should be considered a complementary form of monitoring which fills a number of gaps. It tells us about unexpected outcomes, encourages a diversity of views, enables broad participation, puts events in context and enables a changing focus on what is important. Nevertheless there is scope for improvement:

- MSC can be fine-tuned by developing methods for incorporating insights into programme planning, eliciting the views of programme critics, participatory analysis of stories en masse and improving the feedback process.

- Evaluation approaches that would complement MSC include those that provide quantitative evidence of the achievement of predetermined outcomes, evidence of 'average' experiences and views of non-participants.

- Further research should focus on: the extent of unexpected changes and negative stories that are reported, and ways of strengthening both the feedback loop and the

96

link between dialogue and programme planning. It might also serve to investigate how to strengthen MSC for use in summative evaluation and combine MSC with deductive approaches.

Source: Davies, R., Dart, J., 2005, 'The 'Most Significant Change' (MSC) Technique: A Guide to Its Use', MandE, Melbourne, Australia
Author: Rick Davies and Jess Dart, MandE News, http://www.mande.co.uk/
Full text available online: http://www.mande.co.uk/docs/MSCGuide.pdf

---

5i. Learning from Capacity Building Practice: Adapting the 'Most Significant Change' (MSC) Approach to Evaluate Capacity Building

Author: R Wrigley
Date: 2006
Size: 37 pages

There is growing recognition of the need to take a multi-stakeholder approach to evaluation, which promotes local ownership and builds capacity for reflection, learning, improved performance and self-determination. This paper from the International NGO Training and Research Centre (INTRAC) reflects on the use of the 'Most Significant Change' (MSC) methodology to evaluate the capacity building services of CABUNGO, a local capacity building support provider in Malawi.

As a story based technique, the MSC approach can help to identify and give value to unintended or unexpected changes. Overall, MSC did provide an effective approach to evaluating the capacity of CABUNGO. Using a story-based approach was useful in helping CABUNGO understand its impact on the organisational capacity of its clients and how its services could be improved. The key advantages of using MSC were its ability to capture and consolidate the different perspectives of stakeholders, to aid understanding and conceptualisation of complex change and to enhance organisational learning. The potential constraints lay in meeting the needs of externally driven evaluation processes and dealing with subjectivity and bias.

In the case of CABUNGO, it was decided to concentrate on the three steps that are fundamental to the MSC process - collecting significant change stories; selecting and analysing the most significant stories; and feeding back to stakeholders the results of the selection process - and one additional step - establishing domains of change. The key findings of the evaluation were that:

- CABUNGO has had significant impacts on the sustainability and effectiveness of the non-governmental organisations (NGOs) and community-based organisations (CBOs) with which it has worked.
- the most significant changes in organisational capacity involved shifts in attitudes, skills, knowledge and behaviour, but changes were also seen in relationships and power dynamics.

97

- of the 23 stories, 21 described shifts or improvements to the relationships within the organisation, and of these, 12 also described improved external relationships with the wider community and donors.
- achieving the impacts described depends on preserving the time, resources and expertise that quality capacity building interventions require.
- capacity building providers like CABUNGO face specific challenges in maintaining both the quality of their practice and their long-term financial sustainability.

Many aspects of the process could be improved upon. Difficulties arose from time constraints for using the MSC methodology:

- More time could have been allocated at the beginning for all those directly involved to gain familiarity with the approach. More creative approaches could have been used.
- Employing a participatory approach to find a consensus on the most appropriate domains could have encouraged a more active interest and engagement in the MSC process and outcomes. The second domain could have been phrased more clearly.
- The process of collecting and verifying data could have benefited from interviewing a wider range of stakeholders to gain different perspectives. The selected stories could also have been verified by visiting the places where the described events took place.
- Having more time could have enabled a more in-depth discussion. Encouraging a broader range of stakeholders to attend the Evaluation Summit might have enriched the process. A more participatory alternative would have been to get participants to do their own categorisation of the 'mini' stories and use this as a basis for analysing the many similarities and differences.

Source: Wrigley, R., 2006, 'Learning from Capacity Building Practice: Adapting the 'Most Significant Change' (MSC) Approach to Evaluate Capacity Building', INTRAC Praxis Paper no. 12, International NGO Training and Research Centre, Oxford
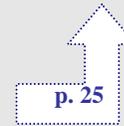**Author:** Rebecca Wrigley , International NGO Training and Research Centre (INTRAC), http://www.intrac.org
Full text available online: http://www.intrac.org/pages/PraxisPaper12.html

5j. Outcome Mapping: Building Learning and Reflection into Development Programs

Author: S Earl et al
Date: 2001
Size: 168 pages

Development organisations are increasingly under pressure to demonstrate that their programmes result in positive changes. This paper from the International Development Research Centre (IDRC) argues that impacts are often the product of events for which no single agency can claim full credit. Outcome mapping moves away from impact assessments to focus on changes in the behaviour of the people it works with directly.

Outcomes are defined as changes in behaviour, relationships, activities, or actions of the people, groups or organisations with whom a programme works directly. Boundary partners are those individuals, groups, and organisations with whom the programme interacts and might influence. Outcome mapping supplements traditional forms of evaluation, which focus on changes in conditions or well-being. It has been developed in organisations where monitoring and evaluation are primarily intended to help with programme learning and improvement.

Outcome mapping differs from traditional forms of evaluation in a number of ways. Some key assumptions and characteristics are:

- linear cause and effect thinking contradicts the understanding of development as a complex process that occurs in open systems. By using outcome mapping, a programme is not claiming achievement of development impacts. The focus is on contributions to outcome, which in turn can enhance the possibility of development impacts.

- outcome mapping focuses on behavioural changes rather than specific impacts. In a cleaner water programme, outcome mapping focuses on whether those responsible for maintaining water purity in communities have the capacity to do so. Traditional evaluation methods would measure the number of filters installed and measure changes in the level of contaminants in the water.

- boundary partners control change. Therefore, as external agents, development programmes only facilitate this process by providing access to new resources, ideas or opportunities.

- outcome mapping is based on principles of participation. It purposefully includes those implementing the programme in the design and data collection so as to encourage ownership and use of findings.

In order to assess whether outcome mapping is appropriate for a programme, the following issues should be considered:

- Strategic directions – Outcome mapping is best used once a programme's strategic directions have been defined. It helps clarify the strategic plan: who is affected, in what ways, and through which programme activities.

- Type of monitoring and evaluation information – Outcome mapping provides information for an evaluation study that examines the programme's performance or the outcomes achieved. It is not a technical evaluation to assess programme relevance or cost-effectiveness.

- Reporting requirements – Outcome mapping depends on self-assessment data generated by the programme team and partners. It requires a commitment to participatory and learning based approaches to learning and evaluation.

- Team consensus – Outcome mapping can provide an opportunity for the group to discuss and negotiate viewpoints systematically and move towards consensus.

- Resource commitments – The programme has to be willing to commit the financial, human and time resources necessary to design and implement a monitoring and evaluation system. A design workshop takes three days; the monitoring system will take one staff member about one day per monitoring session.

Source: Earl S. et al., 2001, 'Outcome Mapping: Building Learning and Reflection into Development Programs', International Development Research Centre, Canada
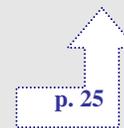**Author:** Sarah Earl, Fred Carden, and Terry Smutylo, IDRC, http://www.idrc.ca

Full text available online: http://www.idrc.ca/en/ev-9330-201-1-DO_TOPIC.html

---

6a. Integrating Gender into the PRS Monitoring and Evaluation

Author: World Bank
Date: 2002
Size: 42 pages

There is growing evidence that gender-sensitive development strategies contribute significantly to economic growth and equity objectives by ensuring that all groups of the poor share in programme benefits. Yet differences between men's and women's needs are often not fully recognised in poverty analysis and participatory planning, and are frequently ignored in the selection and design of Poverty Reduction Strategies (PRSs). A full understanding of the gender dimensions of poverty can significantly change the definition of priority policy and programme interventions supported by the PRS. This chapter from the World Bank PRSP Sourcebook provides practical guidance on identifying and implementing country-level policies and programmes that will benefit both men and women, and maximise potential benefits for poor families.

A PRS involves policies and programme interventions to help the poor across four dimensions of poverty: opportunities (access to labour markets, productive resources, mobility constraints and time burdens), capabilities (access to public services), security (economic vulnerability, civil/domestic violence) and empowerment (political and community voice). In each of these areas, gender must enter into the analysis:

- Opportunities: Gender inequalities in access to credit/financial services include women's limited ownership of land and collateral and reduced business networks.

100

As a result, average loan sizes are smaller than men's. Time constraints on women are often severe, exacerbated by their 'double workday' of production, reproduction and household/community management roles.

- Capabilities: Social stereotyping and family/community socialisation increase the likelihood of girls from poor households leaving education sooner or training for low-income jobs. Access to health services for women is poor, evidenced particularly in the enormous gender differential in Africa's sexual and reproductive burden of disease.

- Security: Women are particularly vulnerable to death, divorce, spousal desertion, community and domestic violence, physical and cultural isolation and marginalisation, ambiguity in legal status/rights, and environmental degradation.

- Empowerment: Women are frequently excluded from the social and political processes affecting them. Gender inequity and powerlessness are culturally ingrained.

Persistent gender inequality in access to and control of a wide range of human, economic, and social assets has direct and indirect costs for economic growth and development and diminishes the effectiveness of poverty reduction efforts. Proactive measures to ensure inclusive participation in the PRS and in the formulation of policies and programmes may take the following forms:

- Addressing gender across the four dimensions of poverty; documenting these experiences; undertaking gender analysis of the data gathered; integrating findings into the country's poverty diagnosis; and defining the policy implications of gender analysis in the country.

- Identifying and integrating gender-responsive priorities into the policy responses and priority actions for the PRS; integrating a gender dimension into the outcome monitoring system and the PRS evaluation strategy; using gender monitoring and impact evaluation results; and building institutional capacity for gender-responsive monitoring and evaluation.

- Including gender-based time budget analysis in the poverty analysis underlying the PRS. One of the highest priorities for a PRS is addressing measures that save time (or improve productivity) and raise labour productivity throughout the economy.

Source: World Bank, 2002, 'Chapter 10: Integrating Gender into the PRS Monitoring and Evaluation', in PRSP Sourcebook, World Bank, Washington, D.C.
Author: World Bank , www.worldbank.org
⬤ Full text available online: http://povlibrary.worldbank.org/files/4221_chap10.pdf

6b. Gender and Monitoring: A Review of Practical Experiences

Author: P Brambilla
Date: 2001
Size: 28 pages

How can monitoring and evaluation (M&E) processes be made gender-sensitive? What measures have organisations taken to assess their effectiveness in mainstreaming gender? This report by BRIDGE for the Swiss Agency for Development and Cooperation (SDC) presents a tool for integrating a gender approach into M&E mechanisms. It draws on a number of case studies and experiences of organisations implementing gender-sensitive M&E mechanisms and provides recommendations.

M&E processes are essential mechanisms to track whether targets are being met and to assess impact. M&E mechanisms can be applied to programmes, projects and policies at field, institutional and government levels. In order to track changes on the situation of women and men and on their relations at all levels, they must be gender-sensitive.

Although many donors and NGOs have started the process of gender mainstreaming, few have developed M&E mechanisms for gender impact. Reviews of NGOs' gender policies generally revealed inadequate M&E mechanisms and gender training, cultural resistance and lack of funding. Other issues included inadequate staffing of women and poor accountability mechanisms. The experiences of some of those NGOs and donor agencies which *have* designed and applied M&E mechanisms from a gender perspective are outlined below:

- The World Bank Toolkit on Water and Sanitation contains numerous instruments for integrating a gender approach into the planning process of water and sanitation. Action Aid has carried out impact assessments in participatory and gender-sensitive ways.
- The International Labour Office (ILO) and the UK Department for International Development (DFID) have each developed guidelines for gender-sensitive M&E mechanisms. The Netherlands Development Organisation (SNV) has developed a manual to carry out participatory gender audits.
- At government level, National Machineries for Gender Equality have been introduced to support government-wide mainstreaming of a gender-equality perspective into all policy areas.
- Further methods to hold governments accountable on gender equality include gender budgets and civil society organisation auditing of government.

On the basis of case studies and different experiences of the implementation of M&E mechanisms, the following recommendations can be made:

- Indicators must be both qualitative and quantitative and include contextual factors. Wherever possible they should measure empowerment issues, such as changes in men's and women's attitudes, perceptions, practices and knowledge.

102

- Indicators do not generally provide information on gender relations, are subject to gender bias and are rarely comparable internationally. Gender-disaggregated indicators are therefore necessary, but must be complemented by qualitative analysis and baseline data in order to track changes in gender relations.
- M&E is needed at every stage of the project, programme or policy implementation cycle. M&E systems must be part of a gender-sensitive planning cycle and have clear objectives against which to measure results and changes.
- M&E processes must be designed in consultation with the target group and carried out in gender sensitive ways.
- Gender training of NGO or government staff and of target groups involved in the M&E process is important. It is also important to extend M&E processes to measure the progress of gender mainstreaming within implementing agencies.

Source: Brambilla, P., 2001, 'Gender and Monitoring: A Review of Practical Experiences', BRIDGE Report 63, BRIDGE, Institute of Development Studies, Brighton
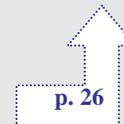**Author:** Paola Bambrilla, BRIDGE, www.bridge.ids.ac.uk
Full text available online: http://www.bridge.ids.ac.uk/reports/re63.pdf

---

6c. Monitoring and Evaluation (M&E) for Development in Peace-Precarious Situation

Author: C Elkins
Date: 2006
Size: 19 pages

How can monitoring and evaluation (M&E) information systems improve programme impact and assist peaceful development in situations prone to violent conflict? This paper from RTI International outlines M&E's status as a unique discipline and describes M&E strategies and tactics implemented in real-world, peace-precarious situations. Even under the stresses of violence and conflict, M&E approaches can help build knowledge of how to push forward peace and development.

M&E supports development by generating relevant, accurate, and timely information. This information enables stakeholders to make evidence-based decisions that can improve programme design and enhance the impact of interventions. Although M&E approaches overlap with academic social science, one major distinction is the typical lack of scientific or quasi-experimental control groups in M&E research. The reasons for this usually relate to budgetary constraints or inadequate human capacity. M&E is necessarily pragmatic. Less precise and less costly data are more acceptable than is the case with academic research. Implementing M&E in peace-precarious situations raises particular challenges, including dramatic changes in the status of institutions and individuals and the physical security of the programme team.

A selection of empirical studies based on the implementation of M&E methods is outlined below:

- In Nepal, Maoist violence disrupted the Maternal and Neonatal Health (MNH) Programme, forcing the M&E plan and survey design to be dynamic. Some initially targeted rural beneficiaries could not be reached, leading to more focus on urban residents.
- Iraq's Local Governance Project (LGP) involved the implementation of quality of life surveys to help stabilise new and emerging local government institutions. The intensity of civil conflict increased data collection costs and the urgency with which results were requested.
- Iraq's Training Model Primary Providers (TMPP) project assisted in the professionalisation of the country's Ministry of Health, especially in assuring the quality of primary health care services. Staff recruitment and retention - local and expatriate - were major pressures.
- In Rwanda, an assessment of health management information systems was complicated by staff reductions in the ministry, a changing political landscape, and the reorganisation of administrative and health districts.
- In Ukraine, the Assocation of Ukrainian Cities (AUC) connects government officials at municipal, city, and village levels across the country. The integration of M&E results information into a website was of critical use for AUC members during a period of intense political uncertainty.

The case studies above show that certain features should be common to all M&E efforts, but may be of particular importance in peace-precarious situations:

- A flexible approach will make use of both quantitative and qualitative data.
- A rapid or rotating calendar of data collecting activities reduces the risks of a fast-changing political context.
- A focus on promoting dialogue and feedback about the information use of M&E will keep projects from becoming out-of-date or irrelevant.
- Integrating the perspective of the 'M&E field' into M&E work will increase awareness of getting data which is 'good enough' where social science purity is impossible.
- A strategic selection of indicators and metrics will focus on what is narrow, redundant, triangular, and may be cautiously interpreted. Such a constrained focus is more likely to produce useful results in difficult circumstances.

Source: Elkins, C., 2006, 'Monitoring and Evaluation (M&E) for Development in Peace-Precarious Situations', RTI International, London
Author: Catherine Elkins, RTI International, http://www.rti.org/
Full text available online: http://www.rti.org/pubs/Elkins_ME_Peace.pdf

6d. Monitoring, Evaluating and Learning for Fragile States and Peacebuilding Programs, Practical Tools for Improving Program Performance and Results

Author: Social Impact
Date: 2006
Size: 168 pages

p. 27

How can organisations implement fragile states peace building (FSP) programmes with realistic development outcomes that can rapidly adapt to changing circumstances? This guide from Social Impact aims to increase the effectiveness of FSP programmes through more systematic approaches to Monitoring, Evaluation and Learning (ME&L). Stronger ME&L enables agencies and communities to understand what is effective, how to consolidate best practice and how to increase accountability to stakeholders.

Monitoring is an assessment of how programmes are progressing towards planned targets, activities and objectives, using mostly quantitative data. It asks 'are we doing the programme right?' Evaluation is a periodic assessment of the relevance, performance, efficiency and impact of a programme. It asks 'are we doing the right programme?' and 'how can we do it better?' ME&L leads to better programmes, deeper learning and increases accountability, whether to donors or the affected people or communities. ME&L in fragile states needs to be swift, safe and highly practical for making management decisions in often unsafe and difficult-to-access environments.

FSP programmes may cover a broad range of objectives. These may range from local governance and legislative reform to confidence building, advocacy, protection of human rights, media strengthening and reintegration of ex-combatants. They require a similarly broad range of ME&L tools that will be practical in very challenging, politically volatile and rapidly changing environments.

- Over past the 10 years there has been a good deal of innovation and learning about performance and evaluation in FSP, with many examples of ME&L good practices emerging from around the world. One example is the Cognitive Social Capital Assessment Tool, which measures trust within and between communities.

- There is a need for ME&L approaches that assess programmes with 'soft' objectives such as trust building. Evaluation approaches for traditional development programmes are inadequate for such objectives. Instead, highly context specific approaches are required to take into account perspectives that are meaningful to the local groups on whom interventions centre.

- Promising innovations have occurred in participatory ME&L (PME&L), which emphasises collaborative analysis, action-learning and empowerment. Such approaches are well suited to FSP, but to be most useful they must be adapted to specific contexts and conflict situations.

Some recommendations are made for using this guide:

- The guide can only offer a snapshot of promising ME&L approaches in a rapidly changing field. Organisations using such approaches must continually innovate

105

and self-reflect so that more is learned about how programmes can become most effective.

- Involving local and community members in ME&L will enable them to become better equipped to manage peace building and transition activities. By working with monitoring experts, local staff and community members can gain critical skills for assessing progress and improving results. This can be part of a legacy that donors and NGOs leave behind.

- The ME&L guide will make the greatest contribution towards better quality programmes if it is used throughout each stage of a programme's life cycle. The guide can strengthen an organisation's strategic plan, programme and project designs, and implementation through better ME&L.

Source: Social Impact, 2006, 'Monitoring, Evaluating and Learning for Fragile States and Peacebuilding Programs, Practical Tools for Improving Program Performance and Results', Social Impact, Arlington, Virginia
Author: Rolf Sartorius and Christopher Carver, Social Impact , www.socialimpact.com
Full text available online:
http://www.socialimpact.com/resource-center/downloads/fragilestates.pdf

---

6e. Nepal Country Assistance Plan: Monitoring in a Fragile State

Author: Department for International Development
Date: 2005
Size: 10 pages

The DFID Nepal Country Assistance Plan (CAP), published in February 2004, aims to reduce poverty and social exclusion and help establish the basis for lasting peace. To this end, CAP commits DFID Nepal to developing improved, locally accountable and transparent systems for monitoring progress. This DFID paper describes the main elements of the CAP Monitoring Framework (CAPMF), setting out how DFID Nepal (DFIDN) proposes to monitor the operating environment for development activities and measure progress towards outputs.

The new CAP Monitoring Framework (CAPMF) is designed to supplement the Nepalese government's Poverty Monitoring and Analysis System (PMAS) launched in May 2004. PMAS - not yet fully operational - will use both input/output and outcome/well-being indicators, gathering data from management information systems in sector ministries, public expenditure tracking, and periodic surveys.

The MF will form a comprehensive picture of programme performance and viability of CAP objectives, addressing the specific monitoring needs for a country in conflict. Complementing conventional project cycle monitoring, CAPMF will strengthen the quality of decision-making at programme and strategic level, providing rigorous evidence of impact on poverty and social exclusion. The MF is also designed to rebalance

DFIDN's effort between programme design/monitoring and between activity/outcome reporting.

CAPMF findings will be reviewed both throughout each year, and annually. An annual report will be disseminated widely amongst Nepal's development community. Strategic direction and priorities for the following year will be agreed and integrated into the programme. The MF comprises three main components:

- Conflict Impact Monitoring: A programme-based monitoring system to track deliverability at an activity level. Quarterly surveys will be analysed for project and staff risk, identifying overall trends across DFIDN's portfolio.
- Livelihoods and Social Inclusion (LSI) Monitoring: DIFDN will produce bi-annual reports analysing the impact of programmes on poor and excluded groups, assessing changes in assets and access to services; voice, influence and claiming rights; the 'rules of the game' (based on socially defined identity) in legal, policy and institutional processes. The bi-annual data will influence future strategic priorities.
- Context Monitoring: Assessing changes in the broader operating environment for development in Nepal, incorporating the resources above with Embassy reports, budgetary trends, fiduciary risk and other sources.

There is a need for continuing and robust analysis of the Nepal context, both in relation to the conflict and also in relation to the wider politics of a state with weak institutions and currently without democratic government. Context analysis is not to become the basis for political conditionality, but to assist DFIDN in adapting aid instruments appropriately in order to achieve greatest impact. A small group has been established to monitor the following specific areas for contextual analysis:

- Maoist tolerance of development activities: monitoring the possibility of increased attempts to control or extort resources from the programme.
- People movement: monitoring the incidence of forced relocation, possible indications of worsening in the conflict signalled by increased migration.
- Additionality and ability to monitor: monitoring donor support, and the resources allocated by government to support the Poverty Reduction Strategy, plus independent monitoring of programme performance on behalf of donors.
- Rule of Law: monitoring the current culture of impunity in relation to human rights abuse.
- Ability to Deliver: monitoring the impact of conflict on community access and security of staff, and how this affects the viability of ongoing programmes.
- Fiduciary risk: monitoring the public sector and tracking progress in anti-corruption and reducing fiduciary risk overall.

Source: Department for International Development, 2005, 'Nepal Country Assistance Plan: Monitoring in a Fragile State', Department for International Development, London
Author: Department for International Development (DFID), http://www.dfid.gov.uk

⬤ Full text available online:
http://www.dfid.gov.uk/pubs/files/monitoring-fragile-state.pdf

Governance and Social Development Resource Centre
International Development Department
University of Birmingham
Edgbaston, Birmingham B15 2TT.
United Kingdom

Tel: 0121 414 5037
Fax: 0121 414 7490

www.gsdrc.org