

# Part 2

## Good practices in country-led monitoring and evaluation systems

Building monitoring and evaluation systems to improve government performance. <i>Keith Mackay, Evaluation Capacity Development Coordinator, Independent Evaluation Group, the World Bank</i> .....	169
Getting the logic right. How a strong theory of change supports programmes which work! <i>Jody Zall Kusek, Lead Coordinator of Global HIV/AIDS Monitoring and Evaluation Group, the World Bank</i> <i>Ray C. Rist, Advisor, the World Bank, and President, International Development Evaluation Association (IDEAS)</i> .....	188
RealWorld Evaluation: conducting evaluations under budget, time, data and political constraints <i>Michael Bamberger, Independent consultant,</i> <i>Jim Rugh, Independent international program evaluator</i> .....	200
Strengthening country data collection systems. The role of the Multiple Indicator Cluster Surveys <i>Marco Segone, Senior Regional Advisor, Monitoring and Evaluation UNICEF CEE/CIS</i> <i>George Sakvarelidze, Monitoring and Evaluation Specialist UNICEF CEE/CIS</i> <i>Daniel Vadnais, Data Dissemination Specialist UNICEF Headquarters</i> .....	238

Strengthening country data dissemination systems.  
Good practices in using DevInfo  
*Nicolas Pron, DevInfo Global Administrator, UNICEF Headquarters*  
*Kris Oswalt, Executive Director, DevInfo Support Group*  
*Marco Segone, Senior Regional Advisor, Monitoring and Evaluation,*  
*UNICEF CEE/CIS*  
*George Sakvarelidze, Monitoring and Evaluation Specialist,*  
*UNICEF CEE/CIS.....252*

Making data meaningful. Writing stories about numbers.  
*UNECE, Statistical dissemination and communication,*  
*Conference of European Statisticians.....268*

# BUILDING MONITORING AND EVALUATION SYSTEMS TO IMPROVE GOVERNMENT PERFORMANCE

*Keith Mackay, Evaluation Capacity Development Coordinator,  
Independent Evaluation Group, the World Bank*

---

## Context

Country-led systems of monitoring and evaluation (M&E) are a concept whose time has come. A growing number of developing and transition countries and most if not all developed countries are devoting considerable attention and effort to their national M&E systems. Many do not label it as such – it may be called evidence-based policy-making, performance-based budgeting, or results-based management, for example – but at the core is an evidentiary system for public sector management that relies on the regular collection of monitoring information and the regular conduct of evaluations.

This paper first examines the various ways in which M&E systems can, and are, used to improve government performance. Key trends influencing developing countries to build or strengthen existing M&E systems are then reviewed. Next, the numerous lessons from international experience in building M&E systems are discussed, including the important role of incentives to conduct and especially to make use of M&E information. Ways to raise awareness of the usefulness of M&E, and to create incentives for the utilization of M&E, are listed. The use of such incentives can help to create demand for M&E. Finally, there is an examination of the importance of conducting a country diagnosis, to provide a shared understanding of the strengths and weaknesses of existing M&E, and, to foster a consensus around an action plan for the further strengthening of M&E.

This paper draws on a recent World Bank book written by the author that discusses all these issues in more depth. The book, *How to build monitoring and evaluation systems to support better government*, is available at:

[http://www.worldbank.org/ieg/ecd/better\\_government.html](http://www.worldbank.org/ieg/ecd/better_government.html)

## Use of monitoring and evaluation systems to improve government performance

M&E can measure the performance of all government policies, programmes, and projects. It can identify what works, what does not, and the reasons why. It also provides information about the performance of individual government ministries and agencies, and of managers and their staff. Additionally, it provides information on the performance of donors who support the work of governments.

The following are four main ways in which monitoring information and evaluation findings can be highly useful to government.

1. To support policy-making, especially budget decision-making (performance-based budgeting) and national planning. These processes focus on government priorities among competing demands from citizens and groups in society. M&E information can support government's deliberations by providing evidence about the most cost-effective types of government activity. Examples of this are different types of employment programmes, health interventions, or conditional cash transfer payments. M&E is widely viewed as a useful tool to help governments under fiscal stress reduce their total spending, by identifying programmes and activities which have relatively low cost-effectiveness. Performance budgeting also helps governments prioritize among competing spending proposals. In this way, it is a vehicle to help them achieve greater value for money from their spending.
2. To help government ministries in their policy development and policy analysis work, and in programme development.
3. To help government ministries and agencies manage activities at the sector, programme, and project levels. This includes government service delivery and the management of staff. M&E identifies the most efficient use of available resources; it can be used to identify implementation difficulties. For example, performance indicators can be used to make cost and performance comparisons (performance benchmarking) among different administrative units, regions, and districts. Comparisons can also be made over time which helps identify good, bad, and promising practices. This can prompt a search for the reasons for this level of performance. Evaluations or reviews are used to identify these reasons. This is the learning function of M&E, and it is often termed "results-based management".

4. To enhance transparency and support accountability relationships by revealing the extent to which government has attained its desired objectives. M&E provides the essential evidence necessary to underpin strong accountability relationships, such as of government to the Parliament or Congress, to civil society, and to donors. M&E also supports the accountability relationships within government, such as between sector ministries and central ministries, and between ministers, managers, and staff. Strong accountability, in turn, can provide powerful incentives to improve performance.

M&E is closely related to many other aspects of public sector management, as listed below.

- Budgetary tracking systems and financial reporting.
- Inter-governmental fiscal relations, including government decentralization, and the extent to which they encompass a focus on government performance.
- Accountability institutions such as national audit offices.
- Commercialization and private sector (profit and nonprofit) delivery of public services, for example, by contracting out government functions. Success in these activities requires a clear understanding of objectives and actual performance.
- Clarification and public reporting of programme goals, objectives, and the strategies necessary for achieving them.
- The setting of explicit customer service standards by service delivery agencies, and monitoring and publicizing the extent to which these are achieved. Civil service reform that focuses on personnel performance management and appraisal, including merit-based hiring, promotion, and firing. This approach recognizes the links between individual performance and project or programme performance.
- The quality of the civil service's policy advice and the extent to which this advice is evidence based (i.e. using M&E).
- Anti-corruption efforts. M&E can be used to identify the "leakage" of government funds by, for example, using public expenditure tracking surveys (PETS). Community monitoring of donor (or government) projects can also be an effective way to help curb corruption in the implementation of projects.

- Finally, M&E provides a vehicle to magnify the voice of civil society, and to put additional pressure on government to achieve higher levels of performance. Civil society (non-government organisations (NGOs), universities, research institutes, think tanks, and the media) can play a role in M&E in several ways, including both as a user and producer of M&E information.

## Key trends influencing developing countries

The example of OECD countries is quite influential in the transition and developing countries. This influence extends to a number of areas of public sector management, such as customer service standards; results-based management; contracting out; privatization; performance pay; decentralization; and, performance budgeting. Most OECD governments place considerable emphasis on the four uses of M&E information: to support evidence-based policy-making (especially performance budgeting); policy development; management; and, accountability. OECD governments collectively possess a great deal of experience in this topic. There is a general understanding that for a government to improve its own performance it needs to devote substantial effort to measuring its performance. As Curristine (2005, pp. 88-89) has noted:

*“Over the past 15 years, the majority of OECD governments have sought to shift the emphasis of budgeting and management away from inputs towards a focus on results, measured in the form of outputs and/or outcomes. While the content, pace, and method of implementation of these reforms varies across countries and over time, they share a renewed focus on measurable results.... In the majority of OECD countries, efforts to assess the performance of programmes and ministries are now an accepted normal part of government. Countries follow a variety of different methods to assess performance, including performance measures, evaluations, and benchmarking.”*

In Latin America, the governments of at least 20 countries are currently working to strengthen their M&E systems. One influence on these governments is the demonstration effect provided by those countries with relatively advanced M&E systems, including Chile; Colombia; Mexico; and, Brazil. Related to this is a common set of economic and social pressures in Latin America. These pressures are the continuing macroeconomic and budgetary constraints; dis-

satisfaction that growth in government spending in the social sectors has not been matched by commensurate increases in the quality and quantity of services provided; continuing pressures to improve and extend government service delivery and income transfers; and, growing pressures for government accountability and for “social control” (i.e. clearer accountability of governments to ordinary citizens and to the congress).

In Eastern Europe an additional influence is seen. Countries which have joined the European Union or are candidate countries are required to strengthen their M&E systems. This is providing further impetus to the trend.

In poorer countries, initiatives of international donors such as the World Bank are also influential. The international debt relief initiative for heavily indebted poor countries has required, as a form of donor conditionality, the preparation of poverty reduction strategy papers (PRSPs) by the countries. These are to include an analysis of each country’s M&E system, in particular, the adequacy of available performance indicators. PRSPs focus on the extent of the country’s success in its poverty-reduction efforts to meet the Millennium Development Goals. However, most poor countries have found it difficult to strengthen their monitoring systems in terms of data production, and especially in terms of data utilization.

At the same time, there are strong accountability pressures on international donors themselves, to demonstrate results from the billions of dollars in aid spent each year, and to place more emphasis on M&E. For the World Bank, for example, these pressures have led to its results agenda. This results agenda requires that the Bank’s country assistance strategies be focused firmly on the extent to which results are actually achieved, and on the Bank’s contribution to them. Another donor trend is a somewhat changing emphasis in the loans made. This change is a move away from narrowly defined projects and toward programmatic lending. This entails provision of block funding, which is, in effect, broad budget support. The absence of clearly defined project activities, and outputs from such lending, also requires a focus on country results, or outcomes, of development assistance. This in turn requires a greater reliance on country systems for national statistics, and for M&E of government programmes.

Donors are working to share their experience, and that of developing countries, in the Managing for Development Results Initiative, which promotes better measurement, monitoring, and manage-

ment for results. This initiative has led to an ambitious programme of activities, including the preparation of a growing collection of resource materials and case studies, from developing countries, concerning the application of M&E and performance management at the national, sector, programme, and project levels.<sup>1</sup>

Multilateral donors who are now heavily engaged in providing support at the country and regional levels to build government M&E systems include the African Development Bank; Asian Development Bank;<sup>2</sup> Inter-American Development Bank; and, the World Bank.<sup>3</sup> A number of bilateral donors are also active in this area. One such is the United Kingdom's Department for International Development (DFID), which has had a particular focus on poverty monitoring systems and the use of performance information to support the budget process.

One final trend influencing the focus on M&E is the growth in the number and membership of national, regional, and global evaluation associations. In Africa, for example, there are now 16 national associations. There are also several regional associations, such as the International Programme Evaluation Network in the Commonwealth of Independent Countries (former Soviet Union countries); the African Evaluation Association; and, in Latin America, Preval and, the new regional association, ReLAC. At the global level there is the International Organisation for Cooperation in Evaluation, and the International Development Evaluation Association. These associations reflect, in part, the growing interest in M&E and the growing number of individuals working in this field. Such communities of practice have the potential to influence the quality of M&E work and thus to facilitate the efforts of governments to strengthen their M&E systems. Some national associations, such as the one in Niger (RenSE), have involved close collaboration among academics, consultants, government officials, and donor officials. This growth has the potential to spread awareness and knowledge of M&E among government officials, and so, to increase demand for it.

- 
- 1 These materials are available at: <http://www.mfdr.org/>
  - 2 [https://wpqp1.adb.org/QuickPlace/cop-mfdr/Main.nsf/h\\_Toc/8d074f8d6f17b0484825712b0028d2fb/?OpenDocument](https://wpqp1.adb.org/QuickPlace/cop-mfdr/Main.nsf/h_Toc/8d074f8d6f17b0484825712b0028d2fb/?OpenDocument)
  - 3 See for example <http://www.worldbank.org/ieg/ecd/>



## Lessons from experience in building monitoring and evaluation systems

There is a growing literature on country experience in building government M&E systems (see, for example, Mackay (2007) and the references there). This literature confirms that there is broad agreement among experts in this area about the key lessons. These are as follows.

1. **Substantive demand from the government is a prerequisite to successful institutionalization.** An M&E system must produce monitoring information and evaluation findings which are judged valuable by key stakeholders; are used to improve government performance; and, which will ensure the funding and continuation of the M&E system. Achieving real demand for M&E is not easy. An important barrier can be a lack of knowledge about what M&E actually encompasses, particularly where the buy-in of key officials is necessary before a lot of effort is put into M&E.

The way around this conundrum is to increase awareness of M&E, in particular, its range of tools, methods, and techniques and, its potential uses. Demand can be increased once key stakeholders in a government begin to understand it better; are exposed to examples of highly cost-effective monitoring systems and evaluation reports; and, when they are made aware of other governments which have set up M&E systems which they value highly. It can also be persuasive to point to the growing evidence of very high returns to investment in M&E.

The supply side is also important including provision of M&E training, manuals, and procedures and the identification of good M&E consultants for example. M&E expertise is certainly necessary if reliable M&E information is to be produced. Those who view M&E in technocratic terms as a stand-alone technical activity tend to focus only on these issues. However, the supply side of producing M&E information is less important than demand. If demand for M&E is strong, then it can be relatively straightforward to improve supply in response, but the converse does not hold.

2. **Incentives are an important part of the demand side.** There need to be strong incentives for M&E to be done well and, in particular, for M&E information to be actually used. Simply having M&E information available does not guarantee use, whether by

programme managers, or by budget officials responsible for advising on spending options, or by a Congress responsible for accountability oversight. This underscores the dangers of a technocratic view which sees M&E as a set of tools with inherent value.

3. **Start with a diagnosis of what M&E functions currently exist** and their strengths and weaknesses, on both the demand and supply sides, when strengthening a government M&E system. The extent of actual utilization of M&E information must be identified, as well as the particular ways in which it is being used. Such diagnoses are themselves a form of evaluation. They are useful for the information and insights they provide, and also because they can be a vehicle for raising the awareness of the importance of M&E and the need to strengthen it.
4. **Find a powerful champion.** This can be a powerful minister or senior official who is able to lead the push to institutionalize M&E; to persuade colleagues about its priority; and, to devote significant resources to create an M&E system. A champion needs to have some understanding of M&E, in terms of tools and methods, and an appreciation of its potential usefulness for government. Government champions have played important roles in the creation of some of the more successful government M&E systems, such as those of Chile, Colombia, and Australia.
5. **Stewardship by a capable ministry.** This related feature of successful government M&E systems is stewardship to drive the design, development, and management of an M&E system. In many developed and upper middle-income countries this has meant the finance ministry. It certainly helps to have the institutional lead of an M&E system close to the center of government, for example, a president's office or a budget office (Bedi and others 2006).

In some countries, capable sector ministries have set up strong M&E systems. A notable example is in Mexico, where the Secretariat for Social Development (SEDESOL), a capable and respected ministry, manages an M&E system that emphasizes both qualitative and impact evaluations. These have included the well-known impact evaluations of the *Progresá* programme. Although expensive, these have been highly influential on the government. The programme now covers some 21 million beneficiaries, and the evaluation can be viewed as having been very cost-effective. Governments in other countries find such

examples of highly influential evaluations to be quite persuasive in relation to the potential usefulness of evaluation, and the merits of setting up a sound M&E system.

The success of M&E in SEDESOL has also helped persuade the powerful finance ministry and the comptroller's office to join the national evaluation council to create a whole-of-government M&E system. This indicates the powerful demonstration effect a successful sector agency can have.

**6. A common mistake is to over-engineer an M&E system.**

This is more readily evident with performance indicators. For example, Colombia's M&E system, SINERGIA, had accumulated 940 performance indicators by 2002. This number was unwieldy for the government's uses of the information for accountability purposes. It has subsequently been reduced to around 500. The appropriate number of performance indicators also depends on the number of government programmes and services and on the type of performance indicator. Senior officials would tend to make use of high-level strategic indicators such as outputs and outcomes. Line managers and their staff, in contrast, would tend to focus on a larger number of operational indicators that target processes and services.

**7. The need to build reliable ministry data systems.** A problem in African countries, and perhaps in some other regions, is that although sector ministries collect a range of performance information, the quality of data is often poor. Data are poor partly because they aren't being used; and they're not used partly because their quality is poor. In such countries there is too much data, not enough information. So, this lesson for the institutionalization of a government M&E system is to build reliable ministry data systems to help provide the raw data on which M&E systems depend. Data verification and credibility is partly a technical issue of accuracy, procedures, and quality control. Related to this issue of technical quality is the need for data to be potentially useful, for it to be available on a timely basis, easy to understand, consistent over time, and so forth.

**8. Utilization is the measure of success of an M&E system.**

The objective of government M&E systems is never to produce large volumes of performance information, or a large number of high-quality evaluations per se. This would reflect a supply-driven approach to an M&E system. Utilization is the measure of success.

9. **Provision of training in a range of M&E tools, methods, approaches, and concepts.** For an M&E system to perform well, it is necessary to have well-trained officials or consultants who are highly skilled in M&E. Thus, most capacity-building plans place considerable emphasis on provision of training in a range of M&E tools, methods, approaches, and concepts. Governments that contract out their evaluations also need to ensure that their officials are able to oversee and manage evaluations. They also need to understand the strengths and limitations (the relative cost-effectiveness) of various types of M&E.
10. **The structural arrangements of an M&E system are important** from a number of perspectives. One is to ensure the objectivity, credibility, and rigor of the M&E information produced by the system. On the data side, governments can rely on external audit committees to verify data. Some rely on the national audit office. Some rely principally on internal ministry audit units. However, some have no audit strategy. On the evaluation side, issues of objectivity and credibility are particularly important. Most Latin American countries deal with this by contracting-out evaluations to external bodies such as academic institutions and consulting firms. This achieves a certain 'distance' between the evaluators and the entities being evaluated, and this has advantages and disadvantages. In contrast, most OECD governments rely on sector ministries to conduct evaluations themselves, although this raises questions about the reliability of self-evaluations.
11. **Building an M&E system is a long-hall effort requiring patience and persistence.** This is the experience of countries that have built a government M&E system. It takes time to create or strengthen data systems; to train or recruit qualified staff; to plan, manage, and conduct evaluations; to build systems for sharing M&E information among relevant ministries; and, to train staff to use M&E information in their day-to-day work, whether that involves programme operations or policy analysis and advice. A handful of countries have been able to create well-functioning evaluation systems (in terms of the quality, number and utilization of the evaluations) within four or five years. In others it has taken more than a decade.
12. **Most countries with well-performing M&E systems have not developed them in a linear manner** according to a set plan. Instead, incremental and even piecemeal approaches seem to be common. One reason for this is the need to make mid-

course corrections as the progress, or lack of progress, with particular M&E initiatives becomes evident. External factors such as a change of government can alter the direction of an M&E system and also, lead to it being significantly strengthened or substantially run down or even abandoned.

13. **The value of regularly evaluating an M&E system.** The frequency of mid-course corrections as M&E systems are being built leads to this additional lesson from experience. Unsurprising, the objective of regular evaluation of the system is to find out what is working, what is not, and why. Such evaluations provide the opportunity to review both the demand and the supply sides of the equation, and to clarify the extent of actual utilization of M&E information, as well as the particular ways in which it is being used.

### **Incentives for conducting and using monitoring and evaluation. How to create demand**

The importance of the demand side has already been noted. However, achieving strong demand within a country is not easy. Having examples of other countries (such as Chile, Colombia, and a number of OECD countries) which have invested the effort necessary to build a well-functioning M&E system, can be enormously influential in creating interest in M&E and building demand for it. Illustrating the cost-effectiveness of individual evaluations conducted in other countries can also persuade decision-makers about the merits of M&E. Some countries, such as Egypt, have developed a good understanding among key government ministers of the potential benefits of M&E. Yet efforts to institutionalize M&E in Egypt have been substantially frustrated by mid-level officials who did not buy into this vision of an M&E system.

The key issue here is the need to ensure there are sufficiently powerful incentives within a government to conduct M&E and to a good quality standard, and to use M&E information intensively. A public sector environment in which it is difficult for managers to perform to high standards and to perform consistently is hostile to M&E. Managers can do little more than focus on narrowly defined day-to-day management tasks. They are not willing to be held accountable for performance if they do not have some surety of the resources available to them or, if they do not have substantial control over the outputs of their activities. In this environment, M&E is understand-

ably seen by managers as probably unfair to them, and as a threat rather than an aid.

The nature of incentives for M&E also depends on how a country envisages using M&E information, whether for the learning function of M&E; or, primarily, for accountability purposes; or, as a tool for performance budgeting; or, if M&E is intended as a tool to support evidence-based policy formulation and analysis. While most countries would claim all these potential uses of M&E information to be important, it is usually the case that one or two predominate. Each of these intended uses of M&E involves different sets of stakeholders and thus incentives to drive the system.

Three types of incentive are presented in Box 1: carrots, sticks, and sermons. Many of these incentives have been used to help institutionalize M&E in developed and developing country governments. Carrots provide positive encouragement and rewards for conducting M&E and utilizing the findings. They include, for example, public recognition or financial incentives to ministries that conduct M&E. Sticks include prods or penalties for ministries or individual civil servants who fail to take performance and M&E seriously. These may include financial penalties for ministries which fail to implement agreed evaluation recommendations. Finally, sermons include high-level statements of endorsement and advocacy concerning the importance of M&E. They also include efforts to raise awareness of M&E and to explain to government officials what's in it for them.

***Box 1: Incentives for conducting and using M&E: carrots, sticks, and sermons***

<b>Carrots</b>	<b>Sticks</b>	<b>Sermons</b>
<ul style="list-style-type: none"> <li>• Awards or prizes – high-level recognition of good or best practice evaluations or of managing for results</li> <li>• Provision of additional funding to ministries to conduct M&amp;E</li> </ul>	<ul style="list-style-type: none"> <li>• Enact laws, decrees, or regulations mandating the planning, conduct, and reporting of M&amp;E</li> <li>• Highlight poor quality evaluation planning, data systems, performance indicators, M&amp;E techniques, M&amp;E reporting</li> </ul>	<ul style="list-style-type: none"> <li>• High-level statements of endorsement by president, ministers, heads of ministries, deputies, and so forth</li> <li>• Awareness-raising seminars/workshops to demystify M&amp;E, provide comfort about its feasibility, and to explain what's in it for participants</li> </ul>

<b>Carrots</b>	<b>Sticks</b>	<b>Sermons</b>
<ul style="list-style-type: none"> <li>• Conduct regular “How Are We Doing?” team meetings (managers and staff) to clarify objectives, review team performance, and identify ways to improve it</li> <li>• Assistance to programme areas in conduct of M&amp;E – via help-desk advice, manuals, free training, etc. This makes it easier (reduces the cost) to do M&amp;E and to use the findings</li> <li>• A government-wide network of officials working on M&amp;E. This helps provide identity and support to evaluators (who often feel isolated within each ministry/entity)</li> <li>• Careful knowledge management of evaluation findings – e.g., providing easily understood executive summaries targeted to key audiences</li> <li>• Provision of budget-related incentives to ministries/agencies to improve performance</li> <li>• Greater management autonomy provided to programmes performing well</li> </ul>	<ul style="list-style-type: none"> <li>• Withhold part of funding from ministries/agencies that fail to conduct M&amp;E</li> <li>• Regularly publish information on all programs’ objectives, outputs, and service quality. Performance comparisons are particularly effective in highlighting good performers and embarrassing poor performers</li> <li>• Highlight adverse M&amp;E information in reports to Parliament/Congress and disseminate widely. This can be politically sensitive and overly embarrassing to government</li> <li>• Set challenging but realistic performance targets – stretch targets – which each ministry, agency, and programme manager is required to meet</li> <li>• Require performance exception reporting where targets not met – requires programme areas to explain poor performance (Colombia)</li> </ul>	<ul style="list-style-type: none"> <li>• Use of actual examples of influential M&amp;E to demonstrate its utility and cost-effectiveness</li> <li>• Piloting of some rapid evaluations and impact evaluations to demonstrate their usefulness</li> <li>• Conferences/seminars on good practice M&amp;E systems in particular ministries and in other countries to demonstrate what M&amp;E systems can produce</li> <li>• Advocacy for government M&amp;E on the part of multilateral and bilateral donors in their loans – this highlights and endorses M&amp;E</li> </ul>

<b>Carrots</b>	<b>Sticks</b>	<b>Sermons</b>
<ul style="list-style-type: none"> <li>• Output- or outcome-based performance triggers in World Bank and other donor loans to governments</li> <li>• Performance contracts / pay for civil servants</li> </ul>	<ul style="list-style-type: none"> <li>• Penalize non-compliance with agreed evaluation recommendations</li> <li>• Involve civil society in M&amp;E of government performance, e.g. using citizen report cards, to stimulate better performance and accountability</li> </ul>	

## The importance of country diagnosis

There is no single best approach to a national or sector M&E system. The particular approach a country should use depends on the actual or intended uses of the information such a system will produce. As discussed earlier, these uses range from assisting resource-allocation decisions in the budget process, to helping prepare national and sector planning, to aiding ongoing management and delivery of government services, to underpinning accountability relationships.

Efforts to build or strengthen government M&E systems clearly need to be tailored to the needs and priorities of each country. Conducting a diagnosis of M&E activities is desirable because it can guide the identification of opportunities for institutionalizing M&E. A formal diagnosis helps identify a country's current strengths and weaknesses in terms of the conduct, quality, and utilization of M&E. Additionally, a diagnosis is invaluable in providing the basis for preparing an action plan. The action plan should be designed according to the desired future uses of monitoring information and evaluation findings.

A diagnosis can be conducted by government or donors, or it may be desirable jointly. The process of conducting a diagnosis provides an opportunity to get important stakeholders within government, particularly senior officials in the key ministries, to focus on the issue of institutionalizing an M&E system. For most if not all developing countries, there will already be a number of M&E activities and systems. But a common challenge is a lack of coordination or



harmonization between them. This can result in significant duplication of effort. A diagnosis that reveals such problems can provide a stimulus to the government to address the problems. By providing a shared understanding of the nature of the problems, it can also help foster a consensus on what is needed to overcome the problem.

In Uganda, for example, the finding that there were 16 M&E sub-systems in existence raised strong concerns among senior officials. Their response led to a decision to create a national, integrated, M&E system to address the problems of harmonization and excessive demands on the suppliers of monitoring information in sector ministries and agencies and at the facility level.

A diagnosis also provides a baseline for measuring a country's progress over time; it is a long-haul effort to build and sustain both demand and supply for M&E. In this environment, it is important to regularly monitor and evaluate the M&E system itself, just as any area of public sector reform should be regularly assessed. Some aspects of an M&E system are amenable to regular monitoring, such as the number of evaluations completed or the extent to which their recommendations are implemented. Other aspects may require more in-depth evaluation from time to time, such as the extent of utilization of M&E information in budget decision making, or the quality of monitoring data. Thus, a diagnosis is a type of evaluation and can identify the degree of progress achieved and any necessary mid-course corrections.

A diagnosis of M&E would be expected to map out a number of key issues as highlighted in Box 2.

### ***Box 2: Key issues for a diagnosis of a government's M&E system***

1. Genesis of the existing M&E system – Role of M&E advocates or champions; key events that created the priority for M&E information (for example, election of reform-oriented government, fiscal crisis).
2. The ministry or agency responsible for managing the M&E system and planning evaluations – Roles and responsibilities of the main parties to the M&E system, for example, finance ministry, planning ministry, president's office, sector ministries, the Parliament or Congress; possible existence of several, uncoordinated M&E systems at the national and sector levels; importance of federal/state/local issues to the M&E system.
3. The public sector environment and whether it makes it easy or difficult for managers to perform to high standards and to be held accountable

for their performance – Incentives for the stakeholders to take M&E seriously, strength of demand for M&E information. Are public sector reforms under way that might benefit from a stronger emphasis on the measurement of government performance, such as a poverty-reduction strategy, performance budgeting, strengthening of policy analysis skills, creation of a performance culture in the civil service, improvements in service delivery such as customer service standards, government decentralization, greater participation by civil society, or an anticorruption strategy?

4. The main aspects of public sector management that the M&E system supports strongly – i) Budget decision making, (ii) national or sector planning, (iii) me management, and (iv) accountability relationships (to the finance ministry, to the president's office, to Parliament, to sector ministries, to civil society).
  - Actual role of M&E information at the various stages of the budget process: such as policy advising and planning, budget decision making, performance review and reporting; possible disconnect between the M&E work of sector ministries and the use of such information in the budget process; any disconnect between the budget process and national planning; opportunities to strengthen the role of M&E in the budget.
  - Extent to which the M&E information commissioned by key stakeholders (for example, the finance ministry) is used by others, such as sector ministries; if not used, barriers to utilization; any solid evidence concerning the extent of utilization by different stakeholders (for example, a diagnostic review or a survey); examples of major evaluations that have been highly influential with the government.
5. Types of M&E tools emphasized in the M&E system: regular performance indicators, rapid reviews or evaluations, performance audits, rigorous, in-depth existence of impact evaluations; scale and cost of each of these types of M&E; manner in which evaluation priorities are set – focused on problem programmes, pilot programmes, high expenditure or -visibility programmes, or on a systematic research agenda to answer questions about programme effectiveness.
6. Who is responsible for collecting performance information and conducting evaluations (for example, ministries themselves or academia or consulting firms); any problems with data quality or reliability or with the quality of evaluations conducted; strengths and weaknesses of local supply of M&E; key capacity constraints and the government's capacity-building priorities.
7. Extent of donor support for M&E in recent years; donor projects that support M&E at whole-of-government, sector, or agency levels – Provision of technical assistance, other capacity building and funding for the conduct of major evaluations, such as rigorous impact evaluations.
8. Conclusions: Overall strengths and weaknesses of the M&E system; its sustainability, in terms of vulnerability to a change in government, for example, how dependent it is on donor funding or other support; current plans for future strengthening of the M&E system.

The purpose of a diagnosis is more than a factual stocktaking. It requires careful judgment concerning the presence or absence of the success factors for building an M&E system. It is therefore important to understand the strength of the government's demand for M&E information and whether there is an influential government champion for M&E.

It is also important to know if there are barriers to building an M&E system, such as lack of genuine demand and ownership; lack of a modern culture of evidence-based decision making and accountability (due, in some countries, to issues of ethics or corruption); lack of evaluation, accounting, or auditing skills; or, poor quality and credibility of financial and other performance information. This understanding naturally leads to the preparation of an action plan to strengthen existing M&E systems or to develop a new system entirely.

Although the preceding issues are largely generic to all countries, it is necessary to adjust the focus according to the nature of the country. Middle-income or upper middle-income countries might well possess a strong evaluation community, centered in universities and research institutes. However the supply of evaluation expertise would be much weaker in many of the poorest countries, for example, those that prepare poverty-reduction strategies. Also, poorer countries are likely to have a strong focus on poverty-monitoring systems, in particular, and are likely to experience much greater difficulties in coping with multiple, unharmonized donor requirements for M&E. Donor pressure is often the primary driver of government efforts to strengthen M&E systems, and the strength of country ownership of these efforts may not be strong.

A question that is often asked is: how long should it take to conduct an M&E diagnosis. There is no simple answer to this question. It all depends on the purposes for which a diagnosis is intended, the range of issues under investigation, and the available time and budget. In some cases a week-long mission to a country has provided a sufficient starting point for a broad understanding of the key issues facing a government interested in strengthening its M&E functions. At the other end of the spectrum is a more formal, detailed, and in-depth evaluation of a government evaluation system, such as the one the Chilean government commissioned the World Bank to undertake. The Chile evaluation involved a team of seven people working for many months.

Other issues may need to be investigated in-depth, such as the quality and credibility of monitoring information and of the sector

information systems which provide this information. Another possible issue is the capacity of universities and other organisations that provide training in M&E. Such training is a common element of action plans to help institutionalize M&E.

Depending on the issues to be addressed in a diagnosis, it might well be necessary to assemble a team of experts with a range of backgrounds. A team might therefore include individuals with expertise in some or all of the following: the management of a government M&E system; performance indicators and systems; statistical systems; evaluation; public sector management reform; and, performance budgeting.

Most diagnoses are neither very rapid nor very time consuming or in-depth; they fall between these two extremes. Nevertheless, a sound diagnosis does require considerable care. The expertise and quality of judgment of those who prepare the diagnosis is crucial.

## **Conclusions**

The focus of this paper is on the key lessons for governments in their efforts to build, strengthen, and fully institutionalize their M&E systems, not as an end in itself but to achieve improved government performance. A consistent message argued here is that the bottom-line measure of “success” of an M&E system is utilization of the information it produces. It is not enough to create a system that produces technically sound performance indicators and evaluations. Utilization depends on the nature and strength of demand for M&E information, and this in turn depends on the incentives to make use of M&E. Some governments in developing countries have a high level of demand for M&E; in others the demand is weak or lukewarm. For these latter countries, there are ways to increase demand by strengthening incentives.

One of the key lessons to incorporate into building an M&E system is the importance of conducting a country diagnosis of M&E. It can provide a sound understanding of M&E activities in the government, the public sector environment and opportunities for using M&E information to support core government functions. Such a diagnosis is an important building block for preparing an action plan. A diagnosis can also be a vehicle for ensuring that key government and donor stakeholders have a shared understanding of the issues and of the importance of strengthening M&E.

## References

- Bamberger, Michael. Mackay, Keith and Ooi, Elaine. (2005), *Influential Evaluations: Detailed Case Studies*. Washington, DC: World Bank.
- Bedi, Tara. Coudouel, Aline. Cox, Marcus . Goldstein, Markus and Thornton, Nigel . (2006), *Beyond the Numbers: Understanding the Institutions for Monitoring Poverty Reduction Strategies*. Washington, DC: World Bank.
- Curristine, Teresa. (2005), *Performance Information in the Budget Process: Results of the OECD 2005 Questionnaire*. OECD Journal on Budgeting 5 (2): 87–131.
- Hatry, Harry P. (2006), *Performance Measurement: Getting Results*, 2nd ed. Washington, DC: The Urban Institute Press.
- Hauge, Arild. (2003), *The Development of Monitoring and Evaluation Capacities to Improve Government Performance in Uganda*. No. 10 of Evaluation Capacity Development Working Paper Series. Washington, DC: World Bank.
- IEG (Independent Evaluation Group). (2004), *Monitoring and Evaluation: Some Tools, Methods and Approaches*, 2nd ed. Washington, DC: World Bank.
- Mackay, Keith. (2007), *How to Build M&E Systems to Support Better Government*. Washington, DC: World Bank.
- Mackay, Keith. (1998), *Evaluation Capacity Development: A Diagnostic Guide and Action Framework*. No. 6 of Evaluation Capacity Development Working Paper Series. Washington, DC: World Bank.
- Mackay, Keith, Gladys Lopez-Acevedo, Fernando Rojas, Aline Coudouel, and others. (2007), *A Diagnosis of Colombia's National M&E System, SINERGIA*. No. 17 of Evaluation Capacity Development Working Paper Series. Washington, DC: World Bank.
- May, Ernesto, David Shand, Keith Mackay, Fernando Rojas, and Jaime Saavedra, eds. (2006), *Towards Institutionalizing Monitoring and Evaluation Systems in Latin America and the Caribbean: Proceedings of a World Bank/Inter-American Development Bank Conference*. Washington, DC: World Bank.
- OECD (Organisation for Economic Co-operation and Development). (2007), *Performance Budgeting in OECD Countries*. Paris: OECD.
- OECD (Organisation for Economic Co-operation and Development). (2005), *Modernizing Government: The Way Forward*. Paris: OECD.
- Ravindra, Adikeshavalu. (2004), *An Assessment of the Impact of Bangalore Citizen Report Cards on the Performance of Public Agencies*. No. 12 of Evaluation Capacity Development Working Paper Series. Washington, DC: World Bank.
- Sandoli, Robert L. (2005), Budgeting for Performance in the United States Using the Program Assessment Rating Tool (PART). Presented at a World Bank-Korea Development Institute conference, Improving the Public Expenditure Management System, December 8.
- Toulemonde, Jacques. (2005), *Incentives, Constraints and Culture-Building as Instruments for the Development of Evaluation Demand*. In Building Effective Evaluation Capacity: Lessons from Practice, ed. Richard Boyle and Donald Lemaire, 153-174. New Brunswick, NJ: Transaction Publishers.

# GETTING THE LOGIC RIGHT. HOW A STRONG THEORY OF CHANGE SUPPORTS PROGRAMMES WHICH WORK!

**Jody Zall Kusek**, *Lead Coordinator of Global HIV/AIDS  
Monitoring and Evaluation Group, the World Bank*

**Ray C. Rist**, *Advisor, the World Bank, and President,  
International Development Evaluation Association (IDEAS)*

---

## Introduction

A vital restaurant area in an urban community, called Ninaville, has been experiencing a recent rash of burglaries. A young couple was even attacked in an adjacent parking garage. Restaurant-goers are also increasingly being harassed on the street by local gangs. As a result, fewer people are frequenting this once-popular eating area. Revenues have plunged and employees are being let go. Over a relatively short time, the area has been transformed from a popular gathering place to one where few venture after dark. Streets are in disrepair, buildings are left vacant, and other fixtures left abandoned. Fortunately, there are funds set aside by the state government for urban renewal in five communities. The Government intends to develop and issue substantial new policies and guidelines for zoning of businesses and residential areas in the State. However, they believed that they need a stronger evidence basis from which to develop the new policy. Thus, they hoped that the five urban renewal projects would serve as pilots to help them understand how to effectively develop the new policy. Ninaville would like to submit a proposal to use the funds to help restore the once-thriving restaurant area. The funds would be made available for three years, with twice yearly reporting on renewal progress in order to maintain funding eligibility.

To achieve the overall goal of restoring security in the restaurant area many questions need to be answered. Are people not coming because they do not feel safe? If so what would make them feel safer? Would hiring more policemen work? Would routing out areas where the gangs congregated be the appropriate thing to do? What about more arrests? Perhaps people are not coming because the res-

restaurant area is no longer on a route for public transportation? What about building a pedestrian mall that would attract other shops and activities for the public? To be a successful candidate for the urban renewal funds, each community would need to develop a strong proposal that described how the funds would be used to achieve key urban renewal goals. Communities were asked to include a programme design, implementation plan, budget and timeline. The city council of Ninaville plans to hold a meeting with all interested stakeholders to identify key concerns, and objectives which they hoped would form the outline of a programme proposal.

## **Thinking through the logic of good programme design**

The first task faced by the city council of Ninaville was to make sure that there was agreement on the nature of the problem. Some people focused on the gangs, and saw the need as to rid the community of these thugs. Others said that while the gangs were important, the real problem was loss of jobs. Others thought that the solution was to bring about economic well-being so that the entire community could benefit. They felt that while once a thriving community, there were many factors besides the gangs and crime that prevented the community from being all it could be. The City Council felt that it was important to outline a set of assumptions that were the likely cause of the recent problems and to identify key risks that had to be managed to achieve renewal of the community.

Ninaville is on the right track. Often referred to as the Programme Logic Model or the Theory of Change approach, a good programme theory is needed to think through the assumptions which will guide an organization (e.g. a community, government, or business), towards the design of effective programme interventions; a strong implementation plan; and, where to best spend resources. A good programme theory provides a strong rationale to: (i) get buy-in from key stakeholders; (ii) expend funds; (iii) suggest achievable outcomes and outputs; and (iv) support scale up of pilot projects to larger and more costly projects and programmes. Ninaville recognizes that in order to compete for one of the five pilots, they have to demonstrate that they are to design and implement a strong programme that will result in positive change. They recognize they need a strong programme theory to demonstrate how the interventions they plan to fund will result in the achievement of their goals.

This discussion fits into the theme of country-led evaluations since to successfully build a strong evaluation culture in developing countries there needs to be an emphasis on how evaluation can help deliver information and analysis which strengthens programme delivery. In short, how evaluation can provide coherent and useful theories of change which countries can deploy as they seek to address the problems they have.

We have identified five questions which need to be answered when thinking through the logic of a programme, or its theory of change. This “CORAL” questionnaire aims to support programme planners in addressing the following:

- C** what is the **concern** or concerns most affecting citizens and other stakeholders?
- O** what is the **outcome** or solution sought? In other words, what would success look like?
- R** what are known or likely **risks** which will stop the programme being successfully implemented?
- A** can key **assumptions** be tested and measured with information readily available to determine what is, or is not, working?
- L** can new programme **logic** and knowledge, gained from implementing programme interventions, be regularly fed back into the programme to revise the design and implementation plan as necessary?

## Can performance frameworks or log frameworks provide the basis for good design and evaluation?

In our 2004 book, “Ten Steps to a Results-Based M&E System”<sup>1</sup>, we identify the ten steps that we believe are necessary to building and using an monitoring and evaluation system to manage to results. In our book, we present a logic model in five parts – inputs, activities, outputs, outcomes, and impacts. We explain how most programme theory is designed from inputs to outputs to impacts. This leaves out any thinking on how to design successful behav-

---

1 The authors summarized the book in an article published in the book “*Bridging the gap: The role of monitoring and evaluation in evidence-based policy making*”. The book is available – free of charge – at: <http://www.unicef.org/ceecis/resources.html>

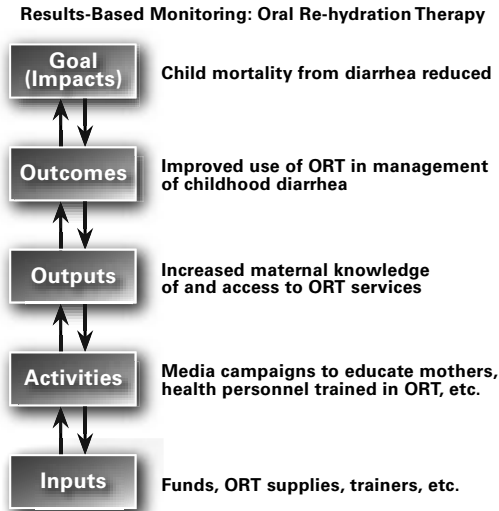


our change and improvements in utilization rates, such as building schools and then actually measuring if children use them rather than theorizing that building 10 new schools will result in improved literacy rates of children. In short, we argue that one cannot get to impacts without first being very clear about what outcomes are to be achieved.

Over the last few years we have heard from numerous programme planners and programme evaluators on the need to further understand what is behind a good performance or logic framework. Questions such as: “how do I know that the interventions in my programme are being designed and implemented to support the programme change I am seeking” or, “how do I keep myself and my staff looking at the big picture”, are frequent. Short of undertaking expensive and often difficult evaluations, it is not easy to know the answers to these questions. However, paying more attention at the design stage will help ensure that a programme will be able to show the effective use of resources, show the links between inputs, activities, outputs and outcomes, and provide a rationale for setting up an evaluation to later test whether the theory “held” or not during implementation. Attention to the programme theory will also help assess, in the case of a programme failure, whether it was the design that failed or whether the implementation failed, or both. Thus a strong programme theory can support effort to better restructure a project to get it back on track.

Figure 1 presents a typical logic model (or results framework as they are often called), for the design of a project intended to support the achievement of reducing mortality rates for children under five years old. Most development programmes are required to include results frameworks to be eligible for international funding. These frameworks intend to demonstrate cause and effect of planned programme components by linking activities and outputs to higher order outcomes and impacts (goals). The suggestion here is that funding media campaigns to inform mothers about the importance of re-hydrating children sick with diarrhoea will ultimately increase their knowledge of its importance and thus change their behaviour towards its use. These activities are presumed causal to the eventual, or higher order, goal of reducing deaths from diarrhoea.

**Figure 1: Example of logic model**



Logic models, or results frameworks, make assumptions that a set of activities are causal to achieving the overall goal. Sometimes these assumptions are made based upon what is considered best practices from similar programmes, or from the findings of evaluation research about what works and why. However, in the rush to get development programmes approved by governments as well as institutional boards, projects are not always designed using valid evidence about what works and why. Assumptions are not tested, and there are no plans to manage risks likely to be encountered during implementation. In these cases, it is down to luck whether the programme theory holds or not.

When the assumptions behind a programme or project design are neither tested nor backed by published evidence, regular “testing” of the logic during implementation can help assure that results will be achieved. This requires that each output and outcome be translated into a set of key performance measures that are tracked regularly to see if the assumptions behind the project or programme are valid. A monitoring system that relies on valid and verifiable information to assess the change of each performance indicator will help determine if the project or programme is achieving planned outputs and outcomes and at what speed. Managers need to pay consistent and regular attention to the original design of the programme and, when necessary, make changes in both the design and the original

assumptions. Building the theory “as you go” requires continued feedback on what appears to be working and what is not and a willingness to make necessary changes to both the original design and assumptions.

In evaluation there is a frequently used phrase, “Weak thrust, weak effect.” This essentially points to the fact that a weakly conceived programme theory of change is not likely to produce strong results, but more likely the opposite: you will not get strong effects from weak designs. Essentially we can think of this in terms of a two by two table (figure 2) showing strong and weak designs across the top and strong and weak implementation along the side. In only one of the four boxes is there both strong theory and strong implementation – which is what it takes for a successful policy or programme or project. Any of the other three boxes represent a problem. Box 2 with a weak design and strong implementation does not provide strong results any more than box 3, with strong design and weak implementation. Finally box 4 is obvious – weak designs and weak implementation can only produce failure. The point of this is that treating design considerations carefully is essential to any opportunity for a successful programme. It cannot happen any other way. A well crafted theory of change is essential for success. Stated differently, both a strong design and strong implementation are requirements if programmes, projects, or policies are to be successful. Neither alone (strong design or strong implementation) is sufficient.

**Figure 2: Weak thrust, weak effect**

		<i>Strength of design</i>	
		Hi	Lo
<i>Strength of implementation</i>	Hi	1	2
	Lo	3	4

## The CORAL questions

Certainly there are many questions that need to be answered during both the design phase of a project and when it is implemented. To assist with this, the authors have, as noted above, developed what we call the CORAL Questionnaire as a self-assessment tool that can be used during the initial design of a new programme or project, during implementation and, to support an evaluation of how well the programme or project achieved its intended goals. In the passages below, we further describe this model.

### ***State the problem that is of concern to key stakeholders***

This is not necessarily self-evident. Different stakeholders can view a problem quite differently, and still all agree there is a problem. The challenge is one of being clear, and in agreement, on the matter of causality. Agreement on the fact that young people are dropping out of school does not automatically lead to agreement on why they are dropping out, let alone what to do about it. The same holds for our example at the beginning of this paper – why is it that the neighbourhood is in decline? Agreement on decline is not hard, but deciding on why it is so can be most contentious. So, to sort out this issue, we need questions such as:

- do we agree there is a problem?
- do we agree on how to define this problem?
- do we agree on the extent of the problem?
- do we agree on what causes this problem?

### ***Agree on desired outcome or solution. Define what success looks like***

If we want to solve our problem, we would have to agree on what a solution would look like. And as our example at the beginning of this paper demonstrates, success can appear very different to different stakeholders. For the owner of the restaurant, it would mean he or she could re-open the restaurant and again make a living; for elderly persons it might mean being able to walk outside without fear of intimidation; for young parents, it might mean being able to again take their children to the playground; and so on. The point is that success is in the eyes of the beholder. But for the evaluator, success is essentially built on the consensus of stakeholders and their view that the theory of change held true; that what was predicted

to take place took place; and, that those who had an input into the discussion on what success would look like, agree that it is what they are seeing. Success is essentially the end point in the theory of change. So, questions that address this issue of success, and what it would look like, might include:

- do we have all the definitions of success on the table?
- do we have ways of measuring success?
- do we agree on when success would be achieved? (Is any neighbourhood ever entirely crime free?)
- can we agree on which definitions of success are most relevant?
- can we articulate the causal model/theory of change on how we will get to that state of success?

### ***Identify and manage risks to success***

There are many factors or risks that can cause success not to happen. Some might be anticipated and we can plan for these; others not (the so-called “unanticipated consequences of social change”). But the fundamental point is that change cannot be completely managed and engineered as one might think could be possible with an infrastructure project. Change takes place within parameters of what are and are not acceptable. A programme might have a trajectory towards success, but it is seldom if ever precisely as was planned or initiated. Multiple circumstances such as clashing personalities of the stakeholders; changes in funding levels; loss of key staff; inability to replace those same staff; and, changes in the political climate, are but a few of a much larger number of threats to the successful completion of the project, programme, or policy. Each of these threats is a risk to the initiative. Each could be enough in the right circumstances to ensure the initiative fails.

The point about identifying and trying to manage risks, is that to ignore them pretty much means one is programming failure. Anticipating how to deal with some of the risks helps boost the prospects of success, but it is not guaranteed that being prepared to mitigate some of the risks will ensure success. The challenge is to think through and acknowledge the key risks, attempt to figure out how to address these risks, and be constantly on the look-out for emergent situations which can sabotage the whole effort. The theories of change for a programme should address the presence of these risks; note how they are going to be addressed; and, establish a monitoring and evaluation system that is flexible, nimble, and

sensitive to information on when things are starting to go wrong. Rigidities in the theory of change are harmful as are rigidities in a monitoring and evaluation system.

Questions to pose here can include:

- have we identified key risks at each stage of the theory of change that threaten the success of the initiative?
- have we decided on how to address these key risks should they emerge?
- do we believe that our monitoring and evaluation system is sufficiently nimble and sensitive to picking up data that show the effort is going off track? (Unanticipated risks are emerging.)

### ***Test key assumptions with valid information***

Assumptions are all those components of a project or programme which are presumed to hold true, to hold constant, or to hold together for the change to eventually occur. Each assumption should be stated explicitly and then examined as to whether it is likely or highly problematic, whether there is research to support it or not, and whether all the key factors, which will facilitate or hinder progress towards the desired change, have been identified within the cumulative total of all assumptions.

A theory of change needs to be continually tested to see if the logic behind it continues to hold during programme or project implementation. To do this, one must ask key questions during design and implementation and when the programme or project is being evaluated.

A theory of change should be able to answer the following:

- is the model an accurate depiction of the programme?
- are all elements and components well defined?
- are there any gaps in the logical chain of events?
- are all elements necessary?
- is the sum total of all elements sufficient to capture the proposed change model?
- are relationships plausible?
- are relationships consistent?
- is the model realistic in terms of the change being achieved?

As described above, we need to regularly test our assumptions by measuring a set of key performance measures designed to track whether desired outputs and outcomes are being achieved. By measuring performance measures on a regular determined basis, managers and decision makers can find out whether projects, programmes and even policies are on track, off track, or even doing better than expected against the targets for performance of those indicators. This provides an opportunity to make adjustments, correct course, and gain valuable institutional and programme experience and knowledge. Ultimately, of course it increases the likelihood of achieving the desired results. In order to test the logic of a programme or project, there must be a valid source of information that can be used to measure each indicator. In accomplishing this, there are nine questions which need to be answered:

- what are the sources of the data that will be used to measure each indicator?
- what are the data collection methods?
- who will collect the data?
- how often will the data be collected?
- what are the costs to collect the data?
- what is the difficulty to collect the data?
- who will analyze the data?
- who will report the data?
- who will use the data?

It should be noted here that no theory of change can be explicit on all possible assumptions. Not all assumptions should be listed and not all assumptions can be tested. The list would be very long -perhaps stretching out with an infinite number of "if-then" statements. As the philosopher E. B. White once noted, "There is no limit to how complicated things can get, on account of one thing always leading to another." What is important is to be relatively sure of getting down with explicit statements all the key assumptions – those presenting the most risk to the programme, whether by happening or by not happening.

## ***Feedback knowledge during implementation to redesign or improve implementation***

Testing key assumptions of the theory of change will produce a continuous flow of information which will support better management of the programme or project, and provide a basis for revising (if necessary) the original design. Thus by allowing flexibility in the programme design logic, decision makers can continuously revise the theory of change if it appears that the original assumptions do not hold. This is not to suggest that poor programme or project performance, due to ineffectual implementation, is a reason for revising the programme logic. If the logic is strong, then the challenge is rightly to improve the implementation – essentially moving from box three to box one in Figure 2.

Key questions which need to be answered, to ensure that knowledge acquired during implementation is used to improve the chances that the programme or project will be successful, include:

- is there a monitoring system in place that allows continual feedback to decision makers?
- is there a champion or individual whose job it is to assess progress towards programme/project implementation?
- are implementation results discussed with key stakeholders?
- what are the opportunities to adjust the original results or performance framework, hence revising the theory of change?

## **Conclusion**

This paper has addressed the issue of why it is important to focus on building coherent logical models so as to be explicit about: (i) what change is anticipated; (ii) what risks there are to that change ever coming into being; (iii) why a system of monitoring is necessary to capture relevant data on whether the change is emerging as planned; and, (iv) how and when relevant stakeholders will be able to decide if the initiative was a success or not. A successful project, programme or policy needs both a strong design and strong implementation. One or other of these two components, by themselves, is not sufficient to ensure success. A well crafted theory of change can help on both accounts, by clearly articulating where the initiative intends to go and, secondly, by matching monitoring data against the theory so as to tell us if the initiative is going in the right direction or not.



## References

- Blamey, A. and Mackenzie, M. (2007). *Theories of Change and Realistic Evaluation*. In: Evaluation. Vol. 13, No. 4.
- Kusek, J. Z. and Rist, R. C. (2008). *Ten Steps to a Results-based Monitoring and Evaluation System*. In: Segone, M., Bridging the gap. The role of monitoring and evaluation in evidence-based policy making. UNICEF
- Kusek, J. Z., Rist, R. C. and White, E. M., (2005). *How will We Know the Millennium Development Goal Results When We See Them?* In: Evaluation. Vol. 11, No. 1.
- Kusek, J.Z. and R. C. Rist, (2004) *Ten Steps to a Results-Based Monitoring and Evaluation System*. Washington, D.C., The World Bank.
- Mason, P. and Barnes, M. (2007). *Constructing Theories of Change*. In: Evaluation. Vol. 13, No. 2.
- Mayne, J. (2007). *Challenges and Lessons in Implementing Results-based Management*. In: Evaluation. Vol 13, no. 1.
- Mayne, J. and Rist, R. C. (2006). *Studies are not Enough: The Necessary Transformation of Evaluation*. In: The Canadian Journal of Evaluation. Vol 21, No. 3.
- Rogers, P. (2008). *Using Program Theory of Evaluate Complicated and Complex Aspects of Interventions*. In: Evaluation. Vol. 14, No. 1.
- Vassen, J. (2006). *Programme Theory Evaluation, Multicriteria Decision Aid and Shareholder Values*. In: Evaluation. Vol. 12, No. 4.
- Zapico-Goni, E. (2007). *Matching Public Management, Accountability, and Evaluation in Uncertain Contexts*. In: Evaluation. Vol. 13, No. 4.

# REALWORLD EVALUATION. CONDUCTING EVALUATIONS UNDER BUDGET, TIME, DATA AND POLITICAL CONSTRAINTS<sup>1</sup>

*Michael Bamberger, Independent consultant,  
Jim Rugh, Independent international program evaluator*

---

## The RealWorld Evaluation context

The RealWorld Evaluation (RWE) approach was developed to assist the many evaluators in both developing, transition, and developed countries, who must conduct evaluations with budget, time, data and political constraints. In one common scenario, the client (project implementing agency, national planning or finance ministry; or, international donor agency), delays contracting an evaluator until late in the project when a decision has to be made on whether to continue support to the project or programme, or possibly to launch a larger second phase. Such tardiness occurs even when evaluation has built into the original project agreement. With the decision point approaching, the funding agency may suddenly realize that it does not have solid information on which to base a decision about future funding of the project; or the project implementing agency may realize it does not have the evidence needed to support its claim that the project is achieving its objectives. An evaluator called in at this point may be told it is essential to conduct the evaluation by a certain date and to produce “rigorous” findings regarding project impact although, unfortunately, very limited funds are available and no systematic baseline data has been collected.

In other scenarios, the evaluator may be called in early in the life of the project but then finds that for budget, political, or methodological reasons, it will not be possible to collect comparison data to determine programme impact by comparing participants with non-

---

1 This article is adapted from the book by Michael Bamberger, Jim Rugh and Linda Mabry. *RealWorld Evaluation: Working under budget, time, data and political constraints* published by Sage in 2006. It also incorporates additional material developed by Bamberger and Rugh for training workshops that have now been offered in 15 countries. Additional materials including more extensive tables are available at [www.realworldevaluation.org](http://www.realworldevaluation.org). The two present authors are entirely responsible for the content and interpretations presented in this chapter.

participants. In some cases, it may not even be possible to collect baseline data on the project participants themselves for purposes of analyzing progress or impact over time. Data constraints may also result from difficulties in collecting information on sensitive topics such as HIV/AIDS; domestic violence; post-conflict reconstruction; or, illegal economic activities (e.g. commercial sex workers, narcotics, or political corruption).

Determining the most appropriate evaluation design under these kinds of circumstances can be a complicated juggling act involving a trade-off between available resources and acceptable standards of evaluation practice. Often the client's concerns are more about budgets and deadlines, and basic principles of evaluation may receive a lower priority. Failure to reach satisfactory resolution of these trade-offs may also contribute to a much lamented problem: low use of evaluation results (see Chelimsky, 1994; Patton, 1997; Operations Evaluation Department, 2004 and 2005). RWE is a response to the all-too-real difficulties in the practical world of evaluation.

The pressures of conducting evaluations under budget and time constraints have often resulted in inattention to sound research design or to identifying and addressing factors affecting the validity of the findings. RWE is based on a seven-step approach, summarized in Figure 1.

## **Scoping the evaluation**

It is important that those charged with conducting an evaluation gain a clear understanding of what those asking for the evaluation (the clients and stakeholders) are expecting – that is, the political setting within which the project and the evaluation will be implemented. It is also important to understand the policy and operational decisions to which the evaluation will contribute and the level of precision required in providing the information which will inform those decisions.

### ***Understanding client's needs***

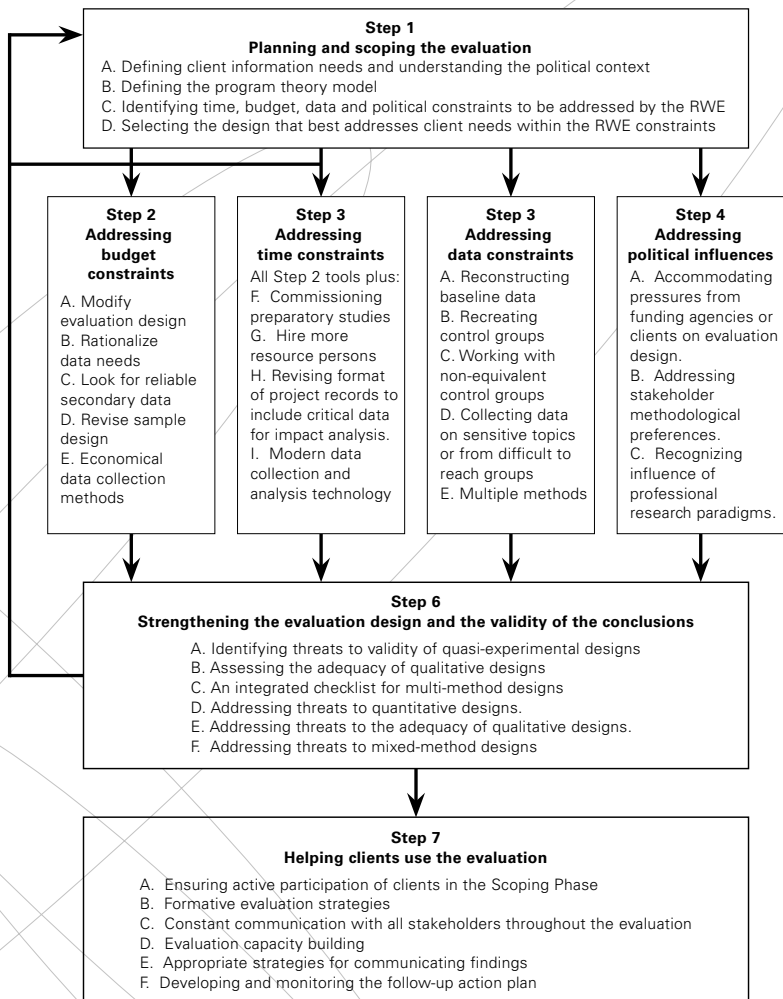
An essential first step in preparing for any evaluation is to obtain a clear understanding of the priorities and information needs of the client (the agency or agencies commissioning the evaluation,) and other key stakeholders (persons interested in or affected by the project). The timing, focus, and level of rigor of the evaluation should be determined by the client information needs and the types of decisions to which the evaluation must contribute.

The process of clarifying what questions need to be answered can help those planning the evaluation to identify ways to eliminate unnecessary data collection and analysis, hence reducing cost and time. The RealWorld evaluator must distinguish between:

- (a) information that is essential to answer the key questions driving the evaluation and,
- (b) additional questions that would be interesting to ask, if there were adequate time and resources, but which may have to be omitted given the limitations faced by the evaluation.

An important function of the scoping phase is to understand whether the lack of consultation with the groups affected by the project (including the poorest and most vulnerable groups), is due to a lack of resources or to the low priority that the client assigns to their involvement. Often, lack of time and money may be used as an excuse, so it is important for the evaluator to fully understand the perspective of the client before deciding what approach to adopt.

**Figure 1: The RealWorld Evaluation [RWE] Approach**



## Understanding the political environment

The political environment includes the priorities and perspectives of the client and other key stakeholders, the dynamics of power and relationships between them and the key players in the project being evaluated, and even the philosophical or methodological biases or preferences of those conducting the evaluation. Table 1 lists some of the ways in which political factors can affect evaluations when they are being designed, while they are being implemented and when the findings are being presented and disseminated.

**Table 1: Examples of some of the ways that political influences affect evaluations**

During evaluation design	
The criteria for selecting evaluators	<p>Evaluators may be selected:</p> <ul style="list-style-type: none"> <li>• for their impartiality or their professional expertise</li> <li>• for their sympathy towards the program</li> <li>• for their known criticisms of the program (in cases where the client wishes to use the evaluation to curtail the program)</li> <li>• for the ease with which they can be controlled</li> <li>• because of their citizenship in the country of the program's funding agency</li> </ul>
The choice of evaluation design and data collection methods	The decision to use either a quantitative or qualitative approach or to collect data that can be put into a certain kind of analytical model (e.g. collecting student achievement or econometric data on an education program) can predetermine what the evaluation will and will not address.
<i>Example of a specific design choice:</i> Whether to use control groups (i.e. experimental or quasi-experimental design)	<p>Control groups may be excluded for political rather than methodological reasons such as:</p> <ul style="list-style-type: none"> <li>• to avoid creating expectations of compensation</li> <li>• to avoid denial of needed benefits to parts of a community</li> <li>• to avoid pressures to expand the project to the control areas</li> <li>• to avoid covering politically sensitive or volatile groups.</li> </ul> <p>On the other hand evaluators may insist on including control groups in the evaluation design in order to follow conventional practice in their profession even when they contribute little to addressing evaluation questions.</p>

The choice of indicators and instruments	The decision to only use quantitative indicators can lead (intentionally or otherwise) to certain kinds of findings and exclude the analysis of other, potentially sensitive topics. For example, issues of domestic violence or sexual harassment on public transport will probably not be mentioned if only structured questionnaires are used.
The choice of stakeholders to involve or consult	The design of the evaluation and the issues addressed may be quite different if only government officials are consulted, compared to an evaluation of the same programme in which community organizations, male and female household heads and NGOs are consulted. The evaluator may be formally or informally discouraged from collecting data from certain sensitive groups, for example by limiting the available time or budget, a subtle way to exclude difficult to reach groups.
Professional orientation of the evaluators	The choice of, for example, economists, sociologists, political scientists or anthropologists to conduct an evaluation will have a major influence on how the evaluation is designed and the findings and recommendations that ensue.
The selection of internal or external evaluators	Evaluations conducted internally by project or agency staff have a different kind of political dynamic and are subject to different political pressures compared to evaluations conducted by external consultants, generally believed to be more independent.  The use of national versus international evaluators also changes the dynamic of the evaluation. For example, while national evaluators are likely to be more familiar with the history and context of the programme, they may be less willing to be critical of programmes administered by their regular clients.
Allocations of budget and time	While budget and time constraints are beyond the total control of some clients, others may try to limit time and resources to discourage addressing certain issues or to preclude thorough, critical analysis.
<b>During implementation</b>	
The changing role of the evaluator	The evaluator may have to negotiate between the roles of guide, publicist, advocate, confidante, hanging judge, and critical friend.

The selection of audiences for progress reports and initial findings	A subtle way for the client to avoid criticism is to exclude potential critics from the distribution list for progress reports. Distribution to managers only, excluding programme staff, or to engineers and architects, excluding social workers and extension agents, will shape the nature of findings and the kinds of feedback to which the evaluation is exposed.
Evolving social dynamics	Often at the start of the evaluation relations are cordial, but they can quickly sour when negative findings begin to emerge or the evaluator does not follow the client's advice on how to conduct the evaluation (e.g. from whom to collect data).
<b>Dissemination and use</b>	
Selection of reviewers	If only people with a stake in the continuation of the project are asked to review the draft evaluation report, the feedback is likely to be more positive than if known critics are involved. Short deadlines, innocent or not, may leave insufficient time for some groups to make any significant comments or to include their comments, introducing a systematic bias against these groups.
Choice of language	In developing countries, few evaluation reports are translated into local languages, thereby excluding significant stakeholders. Budget is usually given as the reason, suggesting that informing stakeholders is not what the client considers valuable and needed. Language is also an issue in the U.S., Canada and Europe where many evaluations concern immigrant populations.
Report distribution	Often, an effective way to avoid criticism is to not share the report with critics. Public interest may be at stake, as when clients have a clear and narrow view of how the evaluation results should be disseminated or used and will not consider other possible uses.

Source: RealWorld Evaluation Table 6.1

It is important to avoid the assumption that political influence is bad and that evaluators should be allowed to conduct the evaluation in the way that they know is "best" without interference from politicians and other "narrow-minded" stakeholders trying to make sure that their concerns are introduced into the evaluation. The whole purpose of evaluation is to contribute to a better understanding of policies and programmes about which people have strong and, often, opposing views. If an evaluation is not subject to any political pressures or influences, this probably means either that the topic being studied is of no consequence to anyone or that the evaluation



is designed in such a way that the concerned groups are not able to express their views. Evaluators should never assume that they are right and that stakeholders who hold different views on the key issues, appropriate methodology, or interpretation of the findings are biased, misinformed, or just plain wrong.

If key groups do not find the analysis credible, then the evaluator may need to go back and check carefully on the methodology and underlying assumptions. It is never an appropriate response to sigh and think how difficult it is to get the client to “understand” the methodology, findings and recommendations.

One of the dimensions of contextual analysis used in developing the programme theory model (see the following section) is to examine the influence of political factors. Many of the contextual dimensions (economic, institutional, environmental, and socio-cultural), influence the way that politically concerned groups will view the project and its evaluation. A full understanding of these contextual factors is essential to understanding the attitudes of key stakeholders to the programme and to its evaluation. Once these concerns are understood, it may become easier to identify ways to address the pressures placed by these stakeholders on the evaluation.

Not surprisingly, many programme evaluations are commissioned with political motives in mind, whether or not they are openly expressed. A client may plan to use the evaluation to bolster support for the programme and may consequently resist the inclusion of anything but positive findings. On the other hand, the real but undisclosed purpose the client may have had for commissioning the evaluation may be to provide ammunition for firing a manager or closing down a project or a department. Seldom, if ever, are such purposes made explicit. Different stakeholders may also hold strongly divergent opinions about a programme, its execution, its motives, its leaders, and how it is to be evaluated. Persons who are opposed to the evaluation being conducted may be able to pre-empt an evaluation or obstruct access to data, acceptance of evaluation results, or continuation of an evaluation contract.

Before the evaluation begins, the evaluator should anticipate these different kinds of potential political issues and try to explore them, directly or indirectly, with the client and key stakeholders.

Political dimensions include not only clients and other stakeholders. They also include individual evaluators, who have preferred approaches that resonate with their personal and professional back-

ground and views as to what constitutes competent, appropriate practice. Different evaluators, even those who have chosen to work together on a project, may take different stances regarding their public and ethical responsibilities. Evaluators, like everyone else, have their own personal values. However, for many evaluators, it may be more comfortable to think of the work of evaluation not as an imposition of the evaluator's values but, rather, as an impartial or objective evidence-based judgment about programme merit, shortcomings, effectiveness, efficiency, and goal achievement. The evaluators must be aware of their own perspectives (and biases) and seek to ensure that these are acknowledged and taken into consideration.

Clients may base their selection of evaluators on their reputations for uncompromising honesty, counting on those reputations to ensure the credibility and acceptance of findings. Or the choice of evaluator may be based on ideological stances the evaluator has taken that are in agreement with the client's. These decisions may be so understated as to initially go unnoticed in friendly negotiations and enthusiastic statements about the strategic importance of the proposed evaluation.

Evaluators should also be alert to the fact that political orientations of clients and stakeholders can influence how evaluation findings are disseminated and used. Clients can sometimes ignore findings they do not like and can suppress distribution by circulating reports only to carefully selected readers, by sharing only abbreviated and softened summaries, and by taking responsibility for presenting reports to boards or funding agencies and then acting on that responsibility in manipulative ways. Clients have been known to give oral presentations and even testimony that distort evaluation findings, to take follow-up activities not suggested by, and even contraindicated by, evaluation reports and, to discredit evaluations and evaluators who threaten their programmes and prestige.

The wise evaluator should be aware of such realities and be prepared to deal with them in appropriate ways during the evaluation design, the implementation of the evaluation and in the presentation and use of the evaluation findings.

## ***Defining the programme theory<sup>2</sup>***

Before an evaluation can be conducted, it is necessary to identify the explicit or implicit theory or logic model that underlies the design upon which a project was based. An important function of an impact evaluation is to test the hypothesis that the project's interventions and outputs contributed to the desired outcomes, which, along with external factors that the project assumed would prevail, were to have led to sustainable impact.

Defining the programme theory or logic model is good practice for any evaluation. It is especially useful in RWE, where, due to budget, time, and other constraints, it is necessary to prioritize what the evaluation needs to focus on. An initial review of what a project did, in the light of its logic model, could reveal missing data or information that is needed to verify whether the logic was sound, and whether the project was able to do what was needed to achieve the desired impact.

If the logic model was clearly articulated in the project plan, it can be used to guide the evaluation. If not, the evaluator needs to construct it based on reviews of project documents and discussions with the project implementing agency, project participants, and other stakeholders. In many cases, this requires an iterative process in which the design of the logic model evolves as more is learned during the course of the evaluation.

In addition to articulating the internal cause-effect theory on which a project was designed, a logic model should also identify the socio-economic characteristics of the affected population groups, as well as contextual factors such as the economic, political, organizational, psychological and environmental conditions which affect the target community.

Every project is designed and implemented within a unique *setting* or *context* that includes local and regional economic, political, institutional, and environmental factors as well as the socio-cultural characteristics of the communities or groups affected by the project. The programme theory must incorporate all these factors through a *contextual analysis*. Where a project is implemented in a number of different locations, it will often be the case that performance and outcomes will differ significantly from one site to another because of the different configurations of contextual variables.

---

2 For a more detailed discussion of program theory models see Bamberger, Rugh and Mabry (2006) RealWorld Evaluation, Chapter 9. This includes references to other recent publications.

## *Customizing plans for evaluation*

Those commissioning an evaluation need to consider a number of factors that should be included in the terms of reference (TOR). The client, and an evaluator (or team of evaluators) being contracted to undertake this assignment, might find the following set of questions helpful to be sure these factors are taken into consideration as plans are made for conducting an evaluation. The answers to these questions can help to focus on important issues to be addressed by the evaluation, including ways to deal with RWE constraints.

- Who asked for the evaluation? Who are the key stakeholders? Do they have preconceived ideas regarding the purpose for the evaluation and expected findings?
- Who should be involved in planning/implementing the evaluation?
- What are the key questions to be answered?
- Will this be a formative or summative evaluation? Is its purpose primarily for learning and improving, accountability, or a combination of both?
- Will there be a next phase, or will other projects be designed based on the findings of this evaluation?
- What decisions will be made in response to the findings of this evaluation? By whom?
- What is the appropriate level of rigor needed to inform those decisions?
- What is the scope/scale of the evaluation?
- How much time will be needed/available?
- What financial resources are needed/available?
- What evaluation design would be required/is possible under the circumstances?
- Should the evaluation rely mainly on quantitative (QUANT) methods, qualitative (QUAL) methods, or a combination of the two?
- Should participatory methods be used?
- Can/should there be a survey of individuals, households, or other entities?
- Who should be interviewed?

- What sample design and size are required/feasible?
- What form of analysis will best answer the key questions?
- Who are the audiences for the report(s)? How will the findings be communicated to each audience?

## **Staffing the evaluation economically**

In this section, we address issues concerning external experts (either from another country or from a different part of the country), content area specialists, and locally available data collectors. The ideal is to compose an evaluation team that includes a good combination of persons with different experiences, skill sets, and perspectives. Where RWE constraints are faced, especially funding, compromises may have to be made in the composition of the evaluation team. Although we address each of these categories of persons separately, it is important to consider the overall combination and the effectiveness of the full evaluation team in meeting the requirements of an evaluation.

### ***Use international consultants wisely***

International consultants are usually contracted:

- because of lack of local technical expertise (inside the organization or in the local research community);
- to build up local capacity;
- to save time;
- to ensure independence and objectivity;
- to ensure expert credibility, and/or;
- because of a requirement by the funding agency.

While, if well selected and used, international consultants can significantly improve the quality of the present and future evaluations, they are also expensive and sometimes disruptive, so they should be selected and used wisely. Under RWE constraints, the goal should be to limit the use of international consultants to those areas where they are essential. Here are a few general rules for selecting and using consultants:

- Ensure that local agencies and the client are actively involved in defining the requirements for the external consultant and in the selection process.

- Consider carefully the relative merits of international and national consultants. There is often a trade-off between greater technical expertise of the international consultant and the local knowledge (and of course language ability) of the national consultant. Not using any national consultants can also antagonize the local professional community who may be reluctant to cooperate with the international expert. It is often a good idea to have an evaluation team that combines the attributes of one or more international evaluators with the right mix of local expertise.
- If an international consultant is used, give priority to candidates who have experience in the particular country and with local language skills (if required).
- For evaluations with an operational focus, avoid selecting consultants with impressive academic credentials but limited field experience in conducting programme evaluations. The purposes and requirements of programme evaluations are different than for academically oriented research.

International consultants are often not used in the most cost-effective way, either because they are doing many things that could be done as well or better by local staff, or because they are brought in at the wrong time. Here are some suggestions on ways to ensure the effective use of international consultants:

- Define carefully what the consultant is being asked to do and consider whether all these activities are necessary.
- Even when the budget is tight, try to plan sufficient time for the consultant to become familiar with the organization, the project, and settings in which it is being implemented. A consultant who does not understand the project, has not spent some time in the communities, or has not built up rapport with project staff, clients, and other stakeholders will be of very little use.
- Plan carefully at what points the consultant should be involved and coordinate ahead of time to ensure that he or she will be available when required. Get tough with consultants who wish to change the timing, particularly at short notice, to suit their own convenience. Some of the critical times to involve a consultant are these:
  - during the scoping phase when critical decisions are being made on objectives, design, and data collection methods and when agreement is being reached with the client on options for addressing time, budget, and data constraints;

- when decisions are being made on sample size and design;
  - when the results of the initial round of data collection are being reviewed and analyzed;
  - when the draft evaluation report is being prepared;
  - when the findings of the evaluation are being presented to the different stakeholders.
- Arrange for a briefing document (preparatory study) to be prepared, by agency staff or local consultants, before the international consultant starts work. This should summarize important information about the project (including compilation of key documents, including monitoring data and periodic reports), key partner agencies, and the settings where the project is located. The document, which should be prepared in coordination with the consultant (for example through an exchange of e-mail or phone calls), might also include rapid diagnostic studies in a few communities. A well-prepared document of this kind can save a great deal of time for the consultant and can initiate dialogue on key issues and priorities among clients, local researchers and stakeholders before the external consultant even arrives.
  - Consider the use of video or phone conferences so that the consultant can maintain more frequent contact with others involved in planning and implementing the evaluation. This enables the consultant to contribute at critical stages of the evaluation without having to always be physically present. In this way, the consultant can make suggestions about the sample or other stages of the design at a sufficiently early stage for it to be possible to make changes based on these recommendations. Video and phone conferences also have the advantage of flexibility, thus avoiding the extremely costly situation where, for example, a consultant flies from Europe to West Africa to participate in the project design phase, only to discover that everything has been delayed for several weeks.

### ***Consider including content area specialists***

In addition to expertise in the relevant evaluation areas (e.g., qualitative interviewing, questionnaire construction, sample design, and data analysis), it is also essential to include at least one team member with the necessary experience in the content area of the evaluation (e.g., agricultural extension, secondary education, micro-credit, health, promoting civil society, etc.). Ideally, if resources permit, the

team should include both a sector expert with experience in many different countries or programmes as well as someone with local knowledge. The school or health system in Chicago or Dushanbe will probably have many unique features (cultural, organizational, and political) which it is important to incorporate into the evaluation.

## Collect data efficiently

### *Simplifying the plans to collect data*

Data collection tends to be one of the most expensive and time-consuming items in an evaluation. Consequently, any efforts to reduce costs or time will almost inevitably involve simplifying plans for data collection. This involves three main approaches (see Table 2):

1. Discuss with the client what information is really required for the evaluation and eliminate other information in the TOR, or mentioned in subsequent discussions, which is not essential in answering the key questions driving this evaluation.
2. Review data collection instruments to eliminate unnecessary information. Data collection instruments tend to grow in length as different people suggest additional items that it would be “interesting” to include, even though not directly related to the purpose of the evaluation.
3. Streamline the process of data collection to reduce costs and time. These include the following:
  - simplifying the evaluation design (e.g. eliminating the collection of baseline data or cutting out the comparison group);
  - clarifying client information needs;
  - look for reliable secondary data;
  - reducing sample size;
  - reducing the costs of data collection, input, and analysis (e.g. use of self-administered questionnaires, using direct observation instead of surveys, using focus groups and community fora instead of household surveys, and finding cheaper data collectors).



### ***Commission preparatory studies***

It is sometimes possible to achieve considerable cost and time savings by commissioning an agency staff person or local consultant to prepare a preparatory study. This can cover these points:

- a description of the different components of the project being evaluated and how they are organized;
- basic information on the implementing agency;
- rapid diagnostic studies of the project communities and possible comparison communities;
- information on government agencies, NGOs and other organizations involved in or familiar with the project;
- recommendations on community leaders and other key informants with whom the international consultant should meet and preparation of background information on them.

### ***Look for reliable secondary data***

A great deal of time and expense can be saved if reliable and relevant secondary data can be obtained. Depending on the country and subjects, it may be possible to find records maintained by government statistical agencies or planning departments; university or other research organizations; schools; commercial banks or credit programmes; mass media; and, many sectors of civil society. Indeed, the evaluator should make use of any relevant records such as monitoring data and annual reports produced by the implementing agency itself.

*Caution: never accept secondary data at face value without checking its reliability and relevance to the communities targeted by the programme being evaluated.*

### ***Collect only the necessary data***

It is important to ensure that only essential information is collected. Long questionnaires and the collection of unnecessary data increases costs and time and also reduces the quality of the information required because respondents become tired if they have to answer large numbers of questions. Therefore, we recommend that all data collection instruments be carefully scrutinized to cut out information that is not relevant and essential to the purpose of the evaluation, and that very likely will never be analyzed or used.

**Table 2: Strategies for addressing data constraints**

<b>Reconstructing Baseline Data</b>		
<i>Approaches</i>	<i>Sources/Methods</i>	<i>Comments/Issues</i>
Using existing documents (secondary data)	<ul style="list-style-type: none"> <li>• Project records</li> <li>• Data from public service agencies (health, education, etc.)</li> <li>• Government household and related surveys</li> </ul>	Consider when the data was collected, what population was included (or excluded), how reliable and relevant the results are in relation to the indicators and population that is being addressed by the present evaluation.
Assessing the reliability and validity of secondary data	<ul style="list-style-type: none"> <li>• School enrollment and attendance records</li> <li>• Patient records in local health centers</li> <li>• Savings and loans cooperatives' records of loans and repayment</li> <li>• Vehicle registrations (to estimate changes in the volume of traffic)</li> <li>• Records of local farmers markets (prices and volume of sales)</li> </ul>	<p>All data must be assessed to determine their adequacy in terms of</p> <ul style="list-style-type: none"> <li>• Reference period</li> <li>• Population coverage</li> <li>• Inclusion of required indicators</li> <li>• Documentation on methodologies used</li> <li>• Completeness</li> <li>• Accuracy</li> <li>• Freedom from bias</li> </ul>

<p>Using recall: asking people to provide numerical (income, crop production, how many hours a day they spent traveling, school fees) or qualitative (the level of violence in the community, the level of consultation of local government officials with the community) at the time the project was beginning</p>	<ul style="list-style-type: none"> <li>• Key informants</li> <li>• PRA (participatory rural appraisal) and other participatory methods</li> </ul>	<p>Recall can be used for</p> <ul style="list-style-type: none"> <li>• School attendance</li> <li>• Sickness/use of health facilities</li> <li>• Income/earnings</li> <li>• Community/individual knowledge and skills</li> <li>• Social cohesion and conflict</li> <li>• Water usage and cost</li> <li>• Major or routine household expenditures</li> <li>• Periods of stress</li> <li>• Travel patterns and transport of produce</li> </ul>
<p>Improving the reliability/validity of recall</p>	<ul style="list-style-type: none"> <li>• Refer to previous research or, where possible, conduct small pretest-posttest studies to compare recall with original information</li> <li>• Identify and try to control for potential bias</li> <li>• Clarify the context</li> <li>• Link recall to important reference points in community or personal history</li> <li>• Triangulation (key informants, secondary sources, PRA)</li> </ul>	<ul style="list-style-type: none"> <li>• Where possible refer to previous research that has determined accuracy of recall on certain types of indicators</li> <li>• Be aware of underestimation of small expenditures, truncating large expenditures by including some expenditures made before the recall period, distortion to conform to accepted behavior, intention to mislead.</li> <li>• Context includes time period, specific types of behavior, reasons for collecting the information</li> </ul>

Key informants	<ul style="list-style-type: none"> <li>• Community leaders</li> <li>• Religious leaders</li> <li>• Teachers</li> <li>• Doctors and nurses</li> <li>• Store owners</li> <li>• Police</li> <li>• Journalists</li> </ul>	Use to triangulate (test for consistency) data from other sources
Collecting sensitive data (e.g., domestic violence, fertility behavior, household decision making and resource control, information from or about women, and information on the physically or mentally handicapped)	<ul style="list-style-type: none"> <li>• Participant observation</li> <li>• Focus groups</li> <li>• Unstructured interviews</li> <li>• Observation</li> <li>• PRA techniques</li> <li>• Case studies</li> <li>• Key informants</li> </ul>	These issues also exist with project participants, but they tend to be more difficult to address with comparison groups because the researcher does not have the same contacts or access to the community.
Collecting data on difficult-to-reach groups (e.g., sex workers, drug or alcohol users, criminals, informal small businesses, squatters and illegal residents, ethnic or religious minorities, and in some cultures, women.)	<ul style="list-style-type: none"> <li>• Observation (participant and non-participant)</li> <li>• Informants from the groups</li> <li>• Self-reporting</li> <li>• Tracer studies and snowball samples</li> <li>• Key informants</li> <li>• Existing documents (secondary data)</li> <li>• Symbols of group identification (clothing, tattoos, graffiti)</li> </ul>	As for previous point

Similarly, the data analysis plan should be reviewed to determine what kinds of disaggregated data analysis are actually required. If it is found that certain kinds of proposed disaggregation are not needed (e.g. comparing the impacts of the project on participants in different locations), then it will often be possible to reduce the size of the sample.

## ***Find simple ways to collect data on sensitive topics and from difficult-to-reach populations***

Another challenge to evaluators, although not unique to RWE, regards the collection of data on sensitive topics such as domestic violence, contraceptive usage, or teenage violence; or from difficult to reach groups such as commercial sex workers, drug users, ethnic minorities, migrants, the homeless, or, in some cultures, women. A number of methods can help to address such topics and reach such groups. However, RWE constraints such as budget, time, or political prejudices could create pressures to ignore these sensitive topics or leave out groups of people who are difficult to reach. There are at least three strategies for addressing sensitive topics:

- Identify a wide range of informants who can provide different perspectives;
- Select a number of culturally appropriate strategies for studying sensitive topics;
- Systematically triangulate;

Difficult-to-reach groups include commercial sex workers, drug or alcohol users, criminals, informal and unregistered small businesses, squatters and illegal residents, ethnic or religious minorities, boy-friends or absent fathers, indentured laborers and slaves, informal water sellers, girls attending boys' schools, migrant workers, and persons with HIV/AIDS, particularly those who have not been tested.

The evaluator may face one of two scenarios. In the first scenario, the groups may be known to exist, but members are difficult to find and reach. In the second scenario, the clients and, at least initially, the evaluator may not even be aware of the existence of such marginalized or "invisible" groups. The techniques for identifying and studying difficult-to-reach groups are similar to those used for addressing sensitive topics and include the following:

- *Participant observation.* This is one of the most common ways to become familiar with and accepted into the milieu where the groups operate or are believed to operate. Often, initial contacts or introductions will be made through friends, family, clients, or in some cases, the official organizations with whom the groups interact.
- *Key informants.* Schedule interviews with persons who are particularly familiar with and well informed about the target groups.

- *Tracer studies.* Neighbors, relatives, friends, work colleagues, and so on are used to help locate people who have moved.
- *Snowball samples.* With this technique, efforts are made to locate a few members of the difficult-to-locate group by whatever means are available. These members are then asked to identify other members of the group so that if the approach is successful, the size of the sample will increase. This technique is often used in the study of sexually transmitted diseases.
- *Socio-metric techniques.* Respondents are asked to identify to whom they go for advice or help on particular topics (e.g., advice on family planning, traditional medicine, or for the purchase of illegal substances). A socio-metric map is then drawn with arrows linking informants to the opinion leaders, informants, or resource persons.

### ***Be creative about data collectors***

Creative options are sometimes available for reducing the cost of contracting data collectors. In a health evaluation, it may be possible to contract student nurses; in an agricultural evaluation, to contract agricultural extension workers; and, for many types of evaluation, to contract graduate students as interviewers or enumerators. Arrangements can often be made with the teaching hospital, the Ministry of Agriculture, or a university professor to contract students or staff at a rate of pay that is satisfactory to them but, well below the market rate. Although these options can be attractive in terms of potential cost savings, or for the opportunity to develop local evaluation capacity, there are obvious dangers from the perspective of quality. The interviewers may not take the assignment very seriously; it may be politically difficult to select only the most promising interviewers; or, to take action against people producing poor-quality work. Supervision and training costs may also be high, and the time required to complete data collection may increase. However, experience shows that these kinds of cooperation can work very well if there is a serious commitment on the part of the agency or university faculty.

Another creative option is to employ data collectors from the community. Sometimes a local high school can conduct a community needs assessment study, or a community organization can conduct baseline studies, or monitor project progress. A number of self-reporting techniques can also be used. For example, individuals or families can keep diaries of income and expenditures, daily time

use, or time, mode, and destination of travel. Community groups can be given cameras, tape recorders, or video cameras and asked to make recordings on issues such as problems facing young people, community needs, or the state of community infrastructure. Although all these techniques pose potential validity questions, they are valuable ways to understand the perspective of the community on the issues being studied.

## **Analyze data efficiently**

### ***Look for ways to manage data efficiently***

Before data can be analyzed, they must be input into an electronic or manual format. If this is not done properly, the quality and reliability of the data can be compromised or time, money, or both can be wasted. Furthermore, if data are not properly managed, there is the risk that significant amounts of information will be lost. The following are some of the main steps in the development and implementation of an analysis plan:

- *Drafting an analysis plan.* This must specify for each proposed type of analysis, the objectives of the analysis, the hypothesis to be tested, the variables included in the analysis, and the types of analysis to be conducted.
- *Developing and testing the codebook.* If there are open-ended questions, the responses must be reviewed to define the categories that will be used. If any of the numerical data have been classified into categories ("More than once a week," "Once a week," etc.), the responses should be reviewed to identify any problems or inconsistencies.
- *Ensuring reliable coding.* This involves both ensuring that the codebook is comprehensive and logically consistent and also monitoring the data-coding process to ensure accuracy and consistency between coders.
- *Reviewing surveys for missing data and deciding how to treat missing data.* In some cases, it will be possible to return to the field or mail the questionnaires back to respondents, but in most cases, this will not be practical. Missing data are often not random, so the treatment of these cases is important to avoid bias. For example, there may be differences between sexes, age, and economic or education groups in their willingness to respond to certain questions. There may also be differences

between ethnic or religious groups or between landowners and squatters. One of the first steps in the analysis should be to prepare frequency distributions of missing data for key variables and, when necessary, to conduct an exploratory analysis to determine whether there are significant differences in missing data rates for the key population groups mentioned above.

*With particular reference to entering the data into the computer or manual data analysis system:*

- *Cleaning the data.* This involves the following:
  - Doing exploratory data analysis to identify missing data and to identify potential problems such as outliers. (These are survey variables where a few scores on a particular variable fall far above or below the normal range.) A few outliers can seriously affect the analysis by making it much more difficult to find statistically significant results (because the standard deviation is dramatically increased). Consequently, the data cleaning process must include clear rules on how to treat outliers.
  - Deciding how to treat missing data and the application of the policies
  - Identifying any variables that may require recoding
  - Providing full documentation of how data were cleaned, how missing data were treated and how any indices were created.

While RWE follows most of the standard data analysis procedures, a number of approaches may be required when time or budget are constraints. When *time* is the main constraint and where additional resources may be available to speed up the process, the following approaches can be considered:

- direct inputting of survey data into hand-held computers;
- use of electronic scanning to read questionnaires;
- sub-contracting data analysis to a university or commercial research organization;
- hiring more, or more experienced, data coders and analysts.

When *money* is the main constraint, one or more of the following options can be considered:

- limiting the kinds of statistical analysis to reduce expensive computer time;



- consider acquiring and using popular statistical packages such as SPSS or SAS so that the analysis can be conducted in-house rather than subcontracting. Needless to say this option requires the availability of statistical expertise in-house.

### ***Focus analysis on answering key questions***

It is wise advice for any evaluation to focus on the key questions that relate to the main purpose of undertaking an assessment. This is especially important for RWE, because choices need to be made on what can be dropped as a consequence of limitations of time and funding. By being reminded of what the major questions are and what is required to adequately answer them, those planning a RWE can be sure to focus on those issues and not others. Typically, the clients and stakeholders, as well as the evaluators themselves, would like to collect additional information. However, when faced with RWE constraints, what would be “interesting to find out” must be culled from “what is essential” to respond to those key questions that drive the evaluation.

The Real-World evaluator must understand which critical issues must be explored in depth and which are less critical and can be studied less intensively or eliminated completely. It is also essential to understand when rigorous (and expensive) statistical analysis is needed by the client (to legitimize the evaluation findings to members of congress or parliament, or to funding agencies critical of the programme), and when more general analysis and findings would be acceptable. The answer to these questions can have a major impact on the evaluation budget and time required, and particularly on the required sample design and size.

## **Assessing and addressing threats to the validity of the evaluation findings and conclusions**

*Validity* refers to the extent to which evaluation findings and conclusions are supported by: the conceptual framework and programme theory model on which the evaluation was based; the statistical techniques (including sample design); how the project was designed and implemented; and, the similarities and differences between the project population and the wider population to which findings are generalized. If there are problems with the evaluation design or the way the data is interpreted, there is a danger that programmes not achieving their intended objectives may be continued

or even expanded, that good programmes may be discontinued or, that priority target groups may not have access to project benefits.

The Appendix to this chapter includes an abbreviated portion of a checklist that has been developed by the authors to assess validity<sup>3</sup>. The checklist<sup>4</sup> identifies seven dimensions of validity and includes indicators for assessing the adequacy with which the evaluation addresses each threat to validity. These are:

- **Objectivity:** Are the conclusions drawn from the available evidence?
- **Reliability:** Is the process of the study consistent, reasonably stable over time and across researchers and methods?
- **Internal validity:** Are the findings credible to the people studied and to readers, and are the presumed causal linkages between project interventions and outcomes valid?
- **Statistical conclusion validity:** The statistical design and analysis may incorrectly assume that programme interventions have contributed to the observed outputs.
- **Construct validity:** The indicators of outputs, impacts and contextual variables may not adequately describe and measure the constructs (hypotheses, concepts) on which the programme theory is based.
- **External validity:** Do the conclusions fit other contexts and how widely can they be generalized?
- **Utilization:** Were findings useful to clients, researchers and communities studied?

The checklist can be used to assess validity at various points in the evaluation:

- (a) When the evaluation design is submitted by the evaluation consultants;
- (b) during the implementation of the evaluation;
- (c) when the draft final evaluation report is submitted;
- (d) After the evaluation has been completed (this is particularly useful for meta-evaluation).

3 The Appendix includes for illustrative purposes the following sections of the checklist: The cover page, the format for the summary assessment of each validity dimension (only two dimensions are included) and examples of the detailed checklists for two dimensions (Objectivity and External Validity)

4 The complete checklist is available at [www.realworldevaluation.org](http://www.realworldevaluation.org).

## Report findings efficiently and effectively

As we mentioned in the section above titled “Customizing Plans for Evaluation”, an evaluation should focus on the key questions which relate to the main reason for the evaluation. This is especially important for RWE, because choices need to be made on what can be dropped because of limitations of time and funding. Those key questions need to be kept in mind not only during the planning for the evaluation, data collection and analysis, but also when the report(s) are being written. There is a temptation to report on all sorts of “interesting findings,” but the evaluator(s) need to keep the report focused on answering the key questions which the client(s) and stakeholders want answered.

One of the most effective ways to increase the likelihood that evaluation findings are used is to ensure that they are of direct practical utility to the different stakeholders.

Some of the factors affecting utilization include:

- timing of the evaluation;
- recognizing that the evaluation is only one of several sources of information and influence on decision makers and ensuring that the evaluation complements these other sources;
- building an ongoing relationship with key stakeholders, listening carefully to their needs, understanding their perception of the political context, and keeping them informed of the progress of the evaluation. There should be “no surprises” when the evaluation report is presented. (Operations Evaluation Department 2005; Patton 1997).

Some steps in the presentation of evaluation findings include the following.

- Understand the evaluation stakeholders and how they like to receive information;
- Use visual presentation to complement written reports or oral presentations. Where appropriate and feasible, make use of presentation tools such as PowerPoint, but do not become a slave to the technology and do be prepared to work without this if the logistics become too complicated. Visual presentations are particularly useful when the presentation is not made in the first language of many people in the audience.

- Share the evaluation results through oral presentations. Many stakeholders are not comfortable with written reports or slide presentations, so talking about the findings can be important.
- Plan the written report to make it simple, attractive, and user-friendly. Consider presenting different versions of the findings in ways that are most understandable and useful to different audiences.
- Involve the mass media. When a goal is to reach and influence a wide audience (e.g. public opinion, all parents of secondary-school-age children, lawmakers), the press can be a valuable ally. However, working with the media requires time and preparation and if their involvement is important, it may be worth hiring a consultant who “knows the ropes.”

### ***Succinct report to primary clients***

The impact of many evaluations is reduced because the findings and recommendations do not reach the primary clients in time and in a form they like and understand. There is no one best way to report evaluation findings. It depends on the clients and the nature of the evaluation. A good starting point is to ask clients which previous reports they found most useful and why.

A general rule, particularly for RWE, where time tends to be a constraint, is to keep the presentation short and succinct. It is a good idea to have a physically short document that can be widely distributed; although the executive summary at the start of a large report may be well written, some clients and stakeholders may be intimidated by the size of the document and may not get round to opening the summary.

Vaughan and Buss (1998) present some useful guidelines for figuring out what to say to busy policy-makers and how to say it. They point out that many policy-makers have the intellectual capacity to read and understand complicated analysis, but most do not have the time. Consequently, many will want to be given a flavor of the complexities of the analysis (they do not wish to be talked down to), but without getting lost in details. Other policymakers may not have the technical background and will want a simpler presentation. So, there is a delicate balance between keeping the respect and interest of the more technical while not losing the less technical. However, everyone is short of time. Therefore the presentation must be short, even if not necessarily simple. Vaughan and Buss’s rules for figuring out what to say are as follows:

- Analyze policy but not politics. Evaluators are hired to provide technical expertise, not to advise on political strategies.
- Keep it simple.
- Communicate reasoning as well as bottom lines. Many policymakers will want to know how the evaluator arrived at the conclusions, so that they can assess how much weight to give to the findings.
- Use numbers sparingly.
- Elucidate, don't advocate. If evaluators advocate particular policies they risk losing the trust of the policymaker.
- Identify winners and losers. Decision makers are concerned with how policies affect their constituencies, particularly in the short run. Consequently, if evaluators and analysts want policymakers to listen to them, they must identify winners and losers.
- Don't overlook unintended consequences. People will often respond to new policies and programmes in unexpected ways, particularly to take advantage of new resources or opportunities. Sometimes unexpected reactions can destroy a potentially good programme, and in other cases unanticipated outcomes may add to the programme's success. Policy-makers are sensitive to the unexpected because they understand the potentially high political or economic costs. Consequently, if the evaluation can identify some important consequences of which policy-makers were not aware, this will catch the attention of the audience and raise the credibility of the evaluation.

### ***Practical, understandable, and useful reports to other audiences***

A dissemination strategy has to be defined to reach groups with different areas of interest, levels of expertise in reading evaluation reports, and preferences in terms of how they like to receive information. In some cases, different groups may also require the report in different languages. The evaluation team must decide which stakeholders are sufficiently important to merit the preparation of a different version of the report (perhaps even translation into a different language) or the organization of separate presentations and discussions.

These issues are particularly important for RWE because reaching the different audiences, particularly the poorest, least educated, and least accessible has significant cost and time implications. There is a danger that when there are budget or time constraints, the evalu-

ation will reach only the primary clients, and many of the groups whose lives are most affected may never see the evaluation, and may never be consulted on the conclusions and recommendations.

An important purpose of the scoping exercise is to agree with the client who will receive and have the opportunity to express opinions about the evaluation report. If the client shows little interest in wider dissemination, but is not actively opposed, then the evaluator can propose cost-effective strategies for reaching a wider audience. If, on the other hand, the client is actively opposed to wider consultation or dissemination, then the evaluator must consider the options – one of which would be to not accept the evaluation contract.

Assuming the main constraints to wider dissemination are time and budget, the following are some of the options:

- Enlist the support of NGOs and civil society organizations. They will often be willing to help disseminate but may wish to present the findings from their own perspective (which might be quite different from the evaluation team's findings), so it is important to get to know different organizations before inviting them to help with dissemination.
- Meetings can be arranged with organizations in the target communities to present the findings and obtain feedback. It is important that these meetings are organized sufficiently early in the report preparation process so that the opinions and additional information can be incorporated into the final report.
- If the client agrees that the findings of the evaluation would be of interest to a broader public, enlist the support of the mass media. It requires certain talents and the investment of a considerable amount of time to cultivate relationships with television, radio, and print journalists. They might be invited to join in field visits or community meetings and they can be sent interesting news stories from time to time. However, working with the mass media can present potential conflicts of interest for the evaluator, and many would argue that this is not an appropriate role for the evaluator.

## **Help clients use the findings well**

Unfortunately, it is all too common for an evaluation to be completed, a formal report written and handed over to the client, and then nothing more done about it. Following the above advice, including involving the client and other key stakeholders throughout the evaluation process, one would hope that the findings of an evalua-

tion are relevant and taken seriously. However, if there is no follow-up, one can be left with the impression that the evaluation had no value. There are examples where major donor agencies, noting the limited use of evaluation reports, have decided to simply stop commissioning routine evaluations. Wouldn't it be better for more effort to be put into making sure evaluations are focused on answering key questions, well done, and then more fully utilized?

A major purpose of RWE is to help those involved focus on what is most important and to be as efficient as possible in conducting evaluations that add value and are useful. The final step, utilization, must be a part of that efficiency formula. If information is not used to inform decisions that lead to improved programme quality and effectiveness, it is wasted. The point here is that those conducting evaluations need to see that the follow-through is an important part of the evaluation process.

One way to do this is to help the client develop an action plan that outlines steps that will be taken in response to the recommendations of an evaluation and then to monitor implementation of that action plan. Doing this is obvious if this was a formative evaluation, where the findings are used to improve subsequent implementation of an ongoing project. Even in the case of a summative evaluation (where the purpose was to estimate the degree to which project outcomes and impacts had been achieved), or where the project that was evaluated has now ended, follow-up should include helping to utilize the lessons learned to inform future strategy and in the design of future projects. At a minimum, those responsible for an evaluation need to do whatever can be done to be sure that the findings and recommendations are documented and communicated in helpful ways to present and future decision makers.

## **Conclusion: who uses RWE, for what purposes and when?**

There are two main users of RWE. These include *evaluation practitioners* who can use the RWE steps and approaches to:

- identify ways to cope with insufficient time and inadequate budgets for evaluations;
- overcome data constraints, particularly the lack of baseline data;
- and identify and address factors affecting the validity and adequacy of the findings of the evaluation.

The other main users are the *clients*, i.e. *representatives of agencies who commission evaluations and/or use evaluation findings*. Their concerns are similar though from different perspectives, including the need to:

- identify ways to reduce the costs of and time for evaluations, while still meeting the requirement for an adequately credible assessment that meets their needs and will be convincing to those to whom they must report; and
- understand the implications of different RWE strategies on the ability of the evaluation to respond to the purposes for which it was commissioned.

Application of the RWE approach can be helpful at three different points in the life of a project or programme: at the start during the planning stage (M&E plan and baseline), when the project is already being implemented (mid-term evaluation) or at the end (final evaluation). When the evaluation planning process begins at the start of the project, RWE can be used to help identify different options for reducing costs or time of the baseline, minimal but relevant monitoring data to be collected throughout the life of the project, plans for the subsequent evaluation(s), and for deciding how to make the best use of available data, or to understand client information needs and the political context within which the evaluation will be conducted.

When the evaluation does not begin until project implementation is already underway, RWE can be used to identify and assess the different evaluation design options that can be used within the budget and time constraints, and to consider ways to reconstruct baseline data. Attention will be given to assessing the strengths and weaknesses of administrative monitoring data available from the project and the availability and quality of secondary data from other sources. The feasibility of identifying a comparison group may also be considered.

When the evaluation does not begin until towards the end of the project (or after the project has already ended), RWE can be used in a similar way to the previous situation except that the design options are more limited as it is no longer possible to observe the project implementation process.

Under any of these scenarios, one of the innovative RWE approaches is to suggest measures that can be taken to strengthen the validity of the findings from the time of initial negotiations of the ToR, during the process of data collection and analysis, and even up to the point when the draft final evaluation report is being reviewed.



## Appendix 1<sup>5</sup>:

### ***CHECKLIST FOR ASSESSING THREATS TO THE VALIDITY OF AN IMPACT EVALUATION<sup>6</sup>***

#### *Part I. Cover Sheet*

<b>1. Name of project/programme</b>
<b>2. Who conducted this validity assessment?</b> (indicate organizational affiliation)
<b>3. When did the evaluation begin?</b> A. Start of the project ____ B. Mid-term ____ C. Towards the end of the project ____ D. When the project has been operating for several years ____
<b>4. At what stage of the evaluation was this assessment conducted?</b> A. Proposed evaluation design ____ B. Progress report on the evaluation ____ C. Draft final evaluation report ____ D. After the evaluation has been completed ____
<b>5. Reason for conducting the threats to validity assessment</b>
<b>6. Summary of findings of the assessment</b>
<b>7. Recommended follow-up actions (if any)</b>

5 The complete checklist is available at [www.realworldevaluation.org](http://www.realworldevaluation.org)

6 Source: Michael Bamberger (2008) adapted from Miles and Huberman (1994) Chapter 10 Section 1; Guba and Lincoln (1989); Shadish, Cook and Campbell (2002) Tables 2.2, 2.4, 3.1 and 3.2; Bamberger, Rugh and Mabry (2006) Chapter 7 and Appendix 1 and Bamberger (2007). The present authors are solely responsible for the adaptation in this abbreviated form.

<b>Part II. SUMMARY ASSESSMENT FOR EACH COMPONENT</b> <i>[see attachments for more detailed assessments]</i>						
	Very strong				Serious problems	Not applicable
	1	2	3	4	5	N/A
<b>Component A. Objectivity</b> ( <i>Confirmability</i> ): <i>Are the conclusions drawn from the available evidence, and is the research relatively free of researcher bias?</i>						
Summary assessment and recommendations						
Overall rating of this component of the evaluation						
Number of issues/problems identified [indicate no. of 4 and 5 ratings]						
<b>Component B. Reliability</b> : <i>Is the process of the study consistent, coherent and reasonably stable over time and across researchers and methods? If emergent designs are used are the processes through which the design evolves clearly documented?</i>						
Summary assessment and recommendations						
Overall rating of this component of the evaluation						
Number of issues/problems identified [indicate no. of 4 and 5 ratings]						

**\*\* Note: This and the following attachment are examples of the detailed checklists that are included for each of the seven components\*\***

<i>Attachment. OBJECTIVITY</i> ( <i>Confirmability</i> )	
<b>Are the conclusions drawn from the available evidence, and is the research relatively free of researcher bias?</b>	<b>Rating</b>
1. Are the conclusions and recommendations presented in the executive summary consistent with, and supported by, the information and findings in the main report.	
2. Are the study's methods and procedures adequately described? Are study data retained and available for re-analysis?	
3. Is data presented to support the conclusions? Is evidence presented to support all findings.	
4. Has the researcher been as explicit and self-aware as possible about personal assumptions, values and biases?	
5. Were the methods used to control for bias adequate?	
6. Were competing hypotheses or rival conclusions considered?	
General comments on this component	
Ratings: 1 = Evaluation design or analysis is very strong; 5 = design or analysis has serious problems	

<i>Attachment. EXTERNAL VALIDITY</i> <i>[Transferability]</i>	
	<b>Rating</b>
Reasons why inferences about how study results would hold over variations in persons, settings, treatments and outcomes may be incorrect.	
<b>1. Sample does not cover the whole population of interest</b> subjects may come from one sex or from certain ethnic or economic groups or they may have certain personality characteristics (e.g. depressed, self-confident). Consequently it may be different to generalize from the study findings to the whole population.	
<b>2. Different settings affect programme outcomes.</b> Treatments may be implemented in different settings which may affect outcomes. If pressure to reduce class size forces schools to construct extra temporary and inadequate classrooms the outcomes may be very different than having smaller classes in suitable classroom settings.	
<b>3. Different outcome measures give different assessments of project effectiveness.</b> Different outcome measures can produce different conclusions on project effectiveness. Micro-credit programmes for women may increase household income and expenditure on children's education but may not increase women's political empowerment.	
<b>4. Programme outcomes vary in different settings.</b> Programme success may be different in rural and urban settings or in different kinds of communities. So it may not be appropriate to generalize findings from one setting to different settings	
<b>5. Programmes operate differently in different settings.</b> programmes may operate in different ways and have different intermediate and final outcomes in different settings. The implementation of community-managed schools may operate very differently and have different outcomes when managed by religious organizations, government agencies and non-governmental organizations.	
<b>6. The attitude of policy makers and politicians to the programme</b> identical programmes will operate differently and have different outcomes in situations where they have the active support of policy makers or politicians than in situations where they face opposition or indifference. When the party in power or the agency head changes it is common to find that support for programmes can vanish or be increased.	

7. <b>Seasonal and other cycles.</b> many projects will operate differently in different seasons, at different stages of the business cycle or according to the terms of trade for key exports and imports. Attempts to generalize findings from pilot programmes must take these cycles into account.	
8. Are the characteristics of the sample of persons, settings, processes, etc. described in enough detail to permit comparisons with other samples?	
9. Does the sample design theoretically permit generalization to other populations?	
10. Does the researcher define the scope and boundaries of reasonable generalization from the study?	
11. Do the findings include enough “thick description” for readers to assess the potential transferability?	
12. Does a range of readers report the findings to be consistent with their own experience?	
13. Do the findings confirm or are they congruent with existing theory? Is the transferable theory made explicit?	
14. Are the processes and findings generic enough to be applicable in other settings?	
15. Have narrative sequences been preserved? Has a general cross-case theory using the sequences been developed?	
16. Does the report suggest settings where the findings could fruitfully be tested further?	
17. Have the findings been replicated in other studies to assess their robustness. If not, could replication efforts be mounted easily?	
<b>General comments on this component</b>	
Ratings: 1 = Evaluation design or analysis is very strong; 5 = design or analysis has serious problems	

## References

- Bamberger, M. (2008) *Enhancing the utilization of evaluation for evidence-based policymaking*, in Segone, Bridging the Gap: the role of monitoring and evaluation in evidence-based policymaking. UNICEF.
- Bamberger, M. (2007) *A framework for assessing the quality, conclusion validity and utility of evaluations. Experience from international development and lessons for developed countries* Paper presented in the Panel Session *Simply the best? Understanding the market for "Good Practice" advice from government research and evaluations* American Evaluation Association Annual Conference. November 2007. Baltimore
- Bamberger, M. and White, H. (2007) *Using Strong Evaluation Designs in Developing Countries: Experience and Challenges* Journal of Multidisciplinary Evaluation. October 2007 Vol 4 No. 8 pp. 58-73 [[http://survey.ate.wmich.edu/jmde/index.php/jmde\\_1/article/view/31/78](http://survey.ate.wmich.edu/jmde/index.php/jmde_1/article/view/31/78)]
- Bamberger, M. (editor). (2000) *Integrating Quantitative and Qualitative Research in Development Projects*. Directions in Development Series. World Bank.
- Brewer, J. & Hunter, A. (2006). (eds) *Foundations of Multimethod Research. Synthesizing Styles*. Sage Publications.
- Brown, J. (2000). *Evaluating the impact of water supply projects in Indonesia*. In Bamberger, M (editor). *Integrating Quantitative and Qualitative Research in Development Projects*. Directions in Development Series. World Bank.
- Creswell, J. C., Clark V.L., Guttman, M.L. & Hanson, W. (2003). *Advanced Mixed Method Research Designs*. In Tashakkori, A., & Teddye, C. (Eds) *Handbook of Mixed Methods in Social and Behavioral Science*. (pp. 209-240). Thousand Oaks, California: Sage Publications.
- Glewwe, P., Kremer, M., Sylvie Moulin & Zitzewitz, E. (2004). *Retrospective vs. prospective analyses of school inputs: the case of flip charts in Kenya*. *Journal of Development Economics* 74: 251-268. <http://www.povertyactionlab.com/projects/project.php?pid=26>
- Guba, E. & Lincoln, Y. (1989) *Fourth Generation Evaluation*. Thousand Oaks, California. Sage Publications.
- Kozel, V. & Parker, B. (2000). *Integrated approaches to poverty assessment in India* in Bamberger, M.(Ed) (2000). *Integrating Quantitative and Qualitative Research in Development Projects. Directions in Development*. (pp. 59-68). Washington D.C: The World Bank.
- Kumar, S. (2002). *Methods for Community Participation. A Complete Guide for Practitioners*. London. ITDG Publishing.
- Miles, M. & Huberman, M. (1994) *Qualitative Data Analysis*. Thousand Oaks, California. Sage Publications.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, California: Sage Publication.

- Patton, M. Q. (1997). (Third Edition). *Utilization-focused evaluation*. Thousand Oaks, California: Sage Publications.
- Ravallion, M. (2006). *Evaluating anti-poverty programmes. Handbook for Agricultural Economics* (edited by Robert Evenson and Paul Schulz) Volume 4. North-Holland
- Rietberger-McCracken, J. & Narayan, D. (1997). *Participatory Rural Appraisal. Module III of Participatory tools and techniques: a resource kit for participation and social assessment*. Environment Department. Washington D.C: The World Bank.
- Roche, C. (1999). *Impact Assessment for Development Agencies. Learning to Value Change*. OXFAM.
- Rugh, J. (1986). *Self-Evaluation: Ideas for Participatory Evaluation of Rural Community Development Projects*. Oklahoma City, Oklahoma: World Neighbors.
- Schwarz, N. & Oyserman, D. (2001). *Asking Questions about Behavior: Cognition, Communication, and Questionnaire Construction*. *American Journal of Evaluation*. Volume 22. No. 2 pp. 127-160.
- Shadish, W. Cook, T. & Campbell, D. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, Houghton Mifflin.
- Tashakkori, A & Teddue, C (eds) (2003). *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, California. Sage Publications.
- Valadez, J. & Bamberger, M. (1994). *Monitoring and evaluating social programmes in developing countries: a handbook for policymakers, managers and researchers*. Washington D.C. World Bank.
- Vaughan, R. & Buss, T. (1994). *Communicating Social Science Research to Policymakers*. Applied Social Science Research Methods Series No. 48. Sage Publications.
- White, H. (2006) *Impact Evaluation: The experience of the Independent Evaluation Group of the World Bank*. Washington D.C.: World Bank ([www.worldbank.org/ieg/ecd](http://www.worldbank.org/ieg/ecd))
- White, H. and Bamberger, M. (2008). *Impact Evaluation in Official Development Agencies*. IDS Bulletin Volume 39 No. 1 March 2008.
- World Bank. Independent Evaluation Group. (2006) *Conducting quality impact evaluations under budget, time and data constraints*. Available free at [www.worldbank.org/ieg/ecd](http://www.worldbank.org/ieg/ecd)
- World Bank. Operations Evaluation Department. (2004). *Influential Evaluations*. World Bank. Operations Evaluation Department. (2005). *Influential Evaluations: Detailed Case Studies*. Available free at [www.worldbank.org/ieg/ecd](http://www.worldbank.org/ieg/ecd).

# STRENGTHENING COUNTRY DATA COLLECTION SYSTEMS. THE ROLE OF THE MULTIPLE INDICATOR CLUSTER SURVEYS

**Marco Segone**, *Senior Regional Advisor, Monitoring and Evaluation, UNICEF CEE/CIS*  
**George Sakvarelidze**, *Monitoring and Evaluation Specialist, UNICEF CEE/CIS*  
**Daniel Vadnais**, *Data Dissemination Specialist, UNICEF Headquarters*

---

## The role of household surveys in country-led monitoring and evaluation systems

Results-based monitoring and evaluation systems are powerful public management tools to demonstrate accountability, transparency and results, as well as to support evidence-based policy making. Good monitoring and evaluation systems need ownership, efficient management, effective maintenance and credibility. The need to strengthen statistical capacity to support the design, monitoring and evaluation of national development plans has been recognized for at least the last three decades. This has been particularly true in the area of monitoring and evaluating of the situation of children and women.

In 1990, for instance, participants of the World Summit for Children recognized that many countries often lack the institutional capacity, or effective systems, for gathering reliable data in a timely manner. UNICEF answered the call and developed the Multiple Indicator Cluster Survey (MICS) programme, with surveys conducted every five years since 1995. Since the initiation of the programme, around 200 surveys have been implemented in approximately 100 countries.

The UNICEF-supported MICS is one of the few household survey programmes that governments can use for collecting standardized information on the socio-economic condition of households and household members, including women and children. Each round of surveys builds upon the last and offers new indicators to monitor current priorities in addition to the monitoring of trends. MICS also



offers a critical look at sub-national disparities faced by particular communities or groups, for instance, the Roma in FYR Macedonia or Serbia.

MICS, along with USAID-supported Demographic and Health Surveys (DHS), provides countries with the opportunities to strengthen their capacity in collecting data that is relevant to national and international development strategies and priorities. Through capacity building activities and a consultative process of adaptation and customization, MICS promotes national ownership of the household survey tool and of the collected data.

### ***Overview of the third round of the Multiple Indicator Cluster Surveys (MICS3)***

The third round of MICS (2005-2007) focused on providing a monitoring tool for the Millennium Development Goals (MDGs) and World Fit for Children Goals, as well as for other major international commitments, such as the United Nations General Assembly Special Session (UNGASS) on HIV/AIDS and the Abuja targets for malaria. Data on nearly half of the MDG indicators were collected in the third round of MICS, offering the largest single source of data for MDG monitoring.

The MICS3 questionnaire collected indicators on a wide range of topics including: child mortality; nutrition; child health; water and sanitation; reproductive health; child development; education; child protection; HIV/AIDS; sexual behaviour; and, children orphaned and made vulnerable by HIV/AIDS.

UNICEF works with a wide range of inter-agency MDG monitoring groups and other inter-agency indicator development groups with the aim of harmonizing, as far as possible, methodologies for measuring priority indicators.<sup>1</sup> UNICEF makes every effort to harmonize MICS – and the indicators measured – with other similar household survey projects, in particular the DHS programme. This level of coordination ensures maximum coverage, analysis of trends over time, and comparability across projects while guaranteeing the acquisition of most of the indicators needed to monitor the situation of children and women locally and globally.

---

1 These groups include: the Inter-agency Group for Child Mortality Estimation, the Malaria monitoring and evaluation reference group, the Technical advisory group of the WHO/UNICEF Joint monitoring programme on water supply and sanitation, the HIV/AIDS Monitoring and evaluation reference group, the Child health epidemiology reference group, the Global Alliance for Vaccines and Immunization Monitoring and evaluation task force and the Countdown to 2015 technical working group.

More than 50 countries carried out MICS3, including 12 countries in Central and Eastern Europe (CEE) and the Commonwealth of Independent States (CIS) which are at the heart of this paper. MICS3 is generating data representative of close to one in four children living in developing countries; nearly two in five children if India and China are excluded<sup>2</sup>. During that round, some 500,000 households were surveyed and more than 300 experts from developing countries were trained in survey methodology.

## Process leading to MICS3 data ownership and use

### *Strengthening national statistical capacity*

Picture 1: First regional MICS3 workshop on Survey planning in Tbilisi, Georgia



The third round of MICS provided a broad avenue for strengthening the national statistical capacity of government institutions and individuals in over 50 countries. A key element of this strategy was UNICEF's implementation of a series

of four regional-level workshops. The purpose of these workshops was to train national officers in charge of implementing MICS3 in their country. Typically, these were government officials representing their national statistical office. For example, in the CEE/CIS region, a total of 12 countries decided to carry out MICS3 and their representatives were invited and trained in the course of the four workshops on household survey planning, data processing, data analysis and report writing and data archiving and dissemination.

The main guidance for MICS3 is available in the Multiple Indicator Cluster Survey Manual 2005, which covers all stages of survey planning and implementation. In addition to the manual, countries that carried out MICS3 were provided with standard software packages, data entry and tabulation programmes, and report templates. Most, but not all countries, followed the guidelines and standard procedures for the implementation of the surveys. UNICEF pro-

2 Source: The State of the World's Children 2008.

vided assistance throughout the survey process, either through the workshops, by distance communications or occasionally by going directly in a country. Throughout this process, all MICS3 participating countries were encouraged to submit to UNICEF key materials such as their national sampling plans, questionnaires, data sets and reports so as to allow the global MICS3 team to review their content and provide feedback.

In 2007-2008, UNICEF commissioned an evaluation of the MICS3 programme. This was carried out by the external consultancy firm John Snow Inc. One component of the evaluation was to assess the guidelines and standard procedures put forward to facilitate the implementation of MICS3. It was found that UNICEF's overall guidance was of high quality and in compliance with current international standards. A vast majority of countries adopted the standard software and data entry and tabulation programmes provided for data processing. This resulted in a significant improvement in standardization of MICS3 data sets. In general, countries that closely followed the MICS3 standards and guidelines and that submitted important materials for review were quite successful in producing data of good quality.

According to the online survey carried out within the framework of the MICS3 Evaluation, 97% of respondents working in implementing agencies felt that the MICS3 helped to build local capacity. The exposure of country level implementation teams to experts; the participation in the regional training workshops; the provision of user-friendly survey guidelines; and, the continuous interaction of the implementation teams with those responsible for the development of tools, have undoubtedly contributed to the development of capacity.

### ***National ownership of MICS3 surveys***

Picture 2: Official signature of the memorandum of understanding between the government of FYR Macedonia and UNICEF



MICS3 promoted the use (or establishment, where not existent) of inter-ministerial steering committees and the development of joint memorandums of understanding. Steering committees included not only government institutions but also international

organizations. They promoted joint review and selection of indicators and modules. This process was part of the assessment of data needs in the countries and allowed for the identification of indicators to fill in the deficit of information for monitoring national strategies, local MDGs and other government priorities.

Picture 3: Local interviewers interviewing the mother of a child in Kazakhstan



The emphasis on national ownership has been a major feature of the MICS programme. In the majority of MICS3 countries, national institutions led all stages of survey planning and implementation. The general approach in MICS3

was to empower national counterparts to undertake all survey activities, and to avoid performing any survey activity on behalf of the country implementers (typically the national statistics offices).

Even when a country required significant amounts of support to carry out a specific survey activity, this was implemented with strong involvement of the government counterparts. The aim was always to leave the completion of the activity to the counterparts. In only a few cases, and only after maximum effort, did UNICEF hire external survey experts to complete the survey, where completion would otherwise have been impossible.

One of the lessons learned from MICS3 is that when government ownership is weak and the national counterparts perceive the survey as a "UNICEF" activity, then the resulting commitment of the implementing agency has also been weak, causing delays in the completion of activities and sometimes sub-standard outputs. Another lesson is that a country's perception of the relevance of MICS has implications for national ownership of the survey and of its results.

## Use of MICS3 data to inform evidence-based policy advocacy

### *Making data meaningful: the importance of data dissemination and communication*

Picture 4: Two-page information sheet on MICS



The newly created dissemination team at UNICEF Headquarters (HQ) has been coordinating a comprehensive global dissemination and communication strategy for MICS data, in close collaboration with MICS3 colleagues in New York, regional and country counterparts. While dissemination materials and tools are country-designed and country-led, the UNICEF HQ team has liaised with MICS3 countries to encourage and support them in planning and delivering a number of

activities. It has also provided technical assistance to many individual countries. As new activities are implemented at the country and regional level, the HQ team has made efforts to track and collect these activities to make them publicly available at [www.childinfo.org](http://www.childinfo.org). These examples have become dissemination models for other countries and regions to use and adapt to their own needs.

To help raise visibility of the MICS tool and increase knowledge about the information it offers, a two-page information sheet on MICS was produced and made available at: [www.childinfo.org](http://www.childinfo.org).

Starting with the planning phase of MICS3, CEE/CIS made special efforts to ensure that MICS findings would be disseminated to the maximum extent possible. CEE/CIS was the first region to host the 4<sup>th</sup> Regional MICS3 Workshops on Data archiving and dissemination, and it actively contributed to making sure one full day would be dedicated to Data dissemination, and one to further analysis. As a result, the third round of MICS saw an increased dissemination of key findings, using new and innovative tools as well as the traditional ones. To access dissemination and further analysis materials based on MICS3 findings from the CEE/CIS region, please visit [http://www.unicef.org/ceecis/resources\\_8588.html](http://www.unicef.org/ceecis/resources_8588.html).

Several countries produced dissemination materials. Serbia and Kyrgyzstan opted for the production of shorter executive versions of MICS3 reports. These are simplified and more user-friendly summaries aiming at conveying the survey messages to the general audience in an efficient manner. Tajikistan designed a calendar highlighting MICS data on a monthly basis; Malawi produced a series of thematic wall charts; Vietnam designed various fact sheets and; Thailand, the first country to have completed MICS3, produced thematic sub-reports and provincial reports, leaflets, fact sheets, and a video.

Almost half of the CEE/CIS countries developed web-pages dedicated to MICS3. Printed materials for dissemination of the survey findings included fact sheets, booklets, leaflets, posters and calendars. Before launching the survey, most countries prepared and distributed media releases which were instrumental to the printing of articles and broadcasting of messages on radio and television.

Picture 5: Press releases were instrumental in producing articles in newspapers highlighting MICS findings



Picture 6: Calendar highlighting MICS findings in Tajikistan



Picture 7: Poster focusing on emerging challenges highlighted by MICS findings in Serbia



In order to make both the process and the content of MICS3 more understandable for the general audience and to promote national ownership of the survey, UNICEF CEE/CIS and HQ supported the development of a comprehensive video on the implementation of MICS3 in Uzbekistan. In addition, Serbia produced 26 episodes of a serial television documentary, called “Serbia fit for children,” based on their MICS findings.

Picture 8: Fact sheet on child nutritional status produced in Tajikistan



Picture 9: Serbia prepared 26 episodes of the TV serial “Serbia fit for children”



To facilitate easy access to MICS3 findings, about 25 countries, including Kyrgyzstan and Tajikistan, created a national version of MICSInfo based on DevInfo - a powerful database system designed to compile and disseminate data. Other countries, including FYR Macedonia and Serbia, included MICS3 data into their existing DevInfo national databases. DevInfo adaptations aim at easier access and dissemination of data on women and children, providing utility for producing charts, tables and maps.

Picture 10: CEE/CIS MICSInfo provides access to key MICS3 findings in 12 countries



The UNICEF CEE/CIS Regional Office produced MICS Info – available at [www.micsinfo.org](http://www.micsinfo.org). It includes MICS3 data from 12 countries disaggregated by: family size; children living arrangement; sex; residence (urban/rural), mother's/caretaker's; wealth index; ethnicity/language/religion.

UNICEF's decision to design a standardized MICS3 final report cover template proved to be very useful by ensuring consistency and a common image among all MICS3 participating countries.

Picture 11: Examples of country adaptations of the MICS final report cover.





Picture 12: New Childinfo website home page



Recently, the UNICEF dissemination team has also made a strong effort to improve the look of the [www.childinfo.org](http://www.childinfo.org) home page which incorporates a number of original features which make it easier for users to find the statistical information they need on children and women.

The website highlights the leading role UNICEF plays in monitoring the situation of children and women worldwide, particularly in terms of: supporting data collection; maintaining and updating global databases; undertaking data analysis and methodological work; promoting data use; and dissemination, as well as being a leader among UN agencies responsible for the global monitoring of the child-related MDGs. The website also provides the technical resources for conducting MICS.

### ***Access to data facilitates further analysis***

MICS3 findings have been instrumental in informing strategic documents produced at global, regional and country level. Further analysis of MICS3 findings has been promoted from the very beginning of the process. One of the major pre-requisites for this was promotion of, and subsequent public access to, the micro datasets through implementing agencies and UNICEF HQ (visit [www.childinfo.org](http://www.childinfo.org)). The International Household Survey Network (IHSN) Microdata Management Toolkit was used to document and archive the data sets and other survey information.

At the global level, an increasing number of analyses (such as a Health Equity study), incorporating MICS3 data, are being carried out. MICS3 data are also the basis for policy analyses in the *Global study on child poverty and disparities*, which is in progress across 40 countries<sup>3</sup>. Country reports, with disaggregated data, are at the heart of the study which will use newly-generated evidence on child poverty from MICS, DHS and other sources, as tools for starting and influencing public policy debates. Study findings will be used to improve access, use, equity and efficacy of social services and benefits, and to strengthen related programmes and partnerships.

3 See the Global Study Guide online at [www.unicefglobalstudy.blogspot.com](http://www.unicefglobalstudy.blogspot.com).

MICS3 data are also being used at the global level by interagency monitoring groups. These groups use MICS findings to develop joint estimates on a number of development indicators, in particular on: child labour; malaria coverage and burden; water and sanitation, immunization; AIDS; and, under-five and infant mortality. A good example is the release of CMEInfo, a DevInfo application presenting child mortality estimates using MICS, DHS and other representative data sources. It is available at: <http://www.childmortality.org/>

MICS3 data have informed a number of key publications, including: *Progress for children: a Report card on maternal mortality*; *Progress on drinking water and sanitation*; *Children and AIDS: Second stock-taking report*; *Countdown to 2015: Tracking progress in maternal, newborn & child survival*<sup>4</sup>; *The State of the world's children: Child survival; malaria and children: progress in intervention coverage*; *Progress for children: a World Fit for Children statistical review*.

Picture 13: Key MICS3 findings from 12 countries are presented in the publication "Emerging challenges for children in Eastern Europe and Central Asia. Focus on disparities"

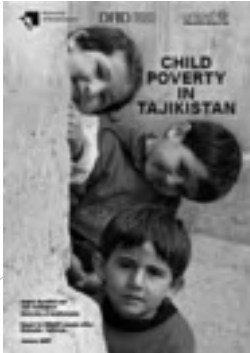


At the regional level, UNICEF CEE/CIS Regional Office used MICS3 data to produce the publication "Emerging challenges for children in Eastern Europe and Central Asia: Focus on disparities". The publication consolidates key findings, focusing on disparities, of 12 MICS3 surveys carried out in CEE/CIS. It comes at a time when there is increasing evidence from a number of sources of growing and disturbing trends towards inequality within countries in the region. The publication presents cross-country tables with data disaggregated by social stratifiers and aims to promote deeper analysis and policy work at country level.

Key regional publications on early childhood development, education and nutrition were also informed by MICS3 data.

4 2015 is the date by which the international community will assess its committed achievement to the MDGs that aim at reducing under-five child deaths by two-thirds, from a baseline set in 1990.

Picture 14: MICS3 data informed the Child poverty study in Tajikistan



Picture 15: MICS3 data informed the study "*The situation of women and children in Serbia. Poor and excluded children*"



Several MICS3 countries, including, in the CEE/CIS region: Albania; Bosnia and Herzegovina; FYR Macedonia; Kyrgyzstan; Serbia; Tajikistan; and, Uzbekistan, used MICS3 data to inform monitoring processes. They use the situation analysis reports related to women and children (including minority groups); child poverty studies; sectoral analysis of early childhood development and child protection; comparative analysis of MICS 2 and MICS3; and monitoring reports for Poverty Reduction Strategies and MDGs.

### ***Use of MICS3 data has enhanced evidence-based policy advocacy and decision making***

MICS3 findings provided participating countries with information disaggregated by several background characteristics such as: region; urban/rural residence; gender; age; level of education; wealth index; ethnicity/language/religion, etc. For many indicators valid data has been obtained on the sub-national level. Disaggregated data allowed for the assessment of disparities within the countries. This is an important aspect for country-led monitoring and evaluation systems. This data also facilitated evidence-based policy advocacy and decision making.

Picture 16: Ms. Ann Veneman, UNICEF's executive director, visiting the MICS stand at the OECD World Forum on measuring and fostering the progress of societies.



When Ms. Ann Veneman, UNICEF's Executive Director, officially revealed (based on new data from MICS, DHS and other reliable sources), that the level of annual deaths of children under the age of five fell, for the first time, below the 10 million mark, news of this child survival

milestone spread all over the world on the Internet, as well as in newspapers, radio and television.

Picture 17: MICS3 findings informed the public hearing at the National Parliament on "Child health. Challenges and Solutions" in Serbia.



At country level, MICS3 findings were presented to Government policy makers and major stakeholders, including to Parliament in Kazakhstan. MICS3 findings have been presented in strategic national Conference, such as at the EU Conference on Social Inclusion

in FYR Macedonia and the National Conference on Poverty in Tajikistan. In Serbia, the MICS3 findings informed the public hearing at the National Parliament on "Child health. Challenges and solutions."

Although still at an early stage, some preliminary results achieved through the use of MICS3 findings in policy making are already being reported. In Serbia, for example, MICS3 findings were instrumental in initiating the establishment of the National commission on young children's' nutrition and feeding practices, as well as the initiative to ban corporal punishment, coordinated by the Serbian NGO network in partnership with the Ministry of Labor and Social Policy.

MICS3 was the first round in which there has been a strong emphasis on dissemination. With materials and activities now available online for countries to use as dissemination models, an increasing number of tools will be developed. This should also ensure that MICS4 data will benefit from an even more elaborate and sophisticated dissemination strategy with the goal of increasing the utilization of the data.

## REFERENCES

John Snow, Inc. (2008). *Evaluation of UNICEF Multiple Indicator Cluster Surveys Round 3 (MICS3)*. Final Report. (forthcoming)

Segone, M. and Sakvarelidze, G. (2008). *Using MICS3 for Evidence-based policy making. The case of CEE/CIS* (PowerPoint presentation).

Available at: [http://www.ceecis.org/mics/MICS3\\_CEECIS\\_DissUse.pdf](http://www.ceecis.org/mics/MICS3_CEECIS_DissUse.pdf)

UNICEF CEE/CIS (2008). *MICS 3 Dissemination: Using MICS3 for Evidence-based policy advocacy. The case of CEE/CIS countries*. Geneva.

Available at: [http://www.unicef.org/ceecis/resources\\_8588.html](http://www.unicef.org/ceecis/resources_8588.html)

UNICEF (2008). *Monitoring the Situation of Children and Women*. New York.

Available at: <http://www.childinfo.org/>

UNICEF (2006). *Multiple Indicator Cluster Survey Manual 2005: Monitoring the Situation of Children and Women*, New York.

Vadnais, Daniel. and Hancioglu, Attila. (2008). *The strategic intent of data collection and analysis. The case of Multiple Indicator Cluster Surveys (MICS)*. In: Segone, M., et al. Bridging the gap. The role of monitoring and evaluation in evidence-based policy making. Switzerland.

# STRENGTHENING COUNTRY DATA DISSEMINATION SYSTEMS. GOOD PRACTICES IN USING DEVINFO

*Nicolas Pron, DevInfo Global Administrator, UNICEF Headquarters*

*Kris Oswald, Executive Director, DevInfo Support Group*

*Marco Segone, Senior Regional Advisor, Monitoring and Evaluation, UNICEF CEE/CIS*

*George Sakvarelidze, Monitoring and Evaluation Specialist, UNICEF CEE/CIS*

---

## Country-led monitoring and evaluation systems are vital to national and decentralized development

Since their adoption by all United Nations Member States in 2000, the Millennium Declaration and the Millennium Development Goals have become a universal framework for development. They are also a means for developing and transition countries, and their development partners, to work together in pursuit of a shared future for all. In 2007, halfway to the MDGs' 2015 target date, there have been gains, but much remains to be done if millions of people are to realize the basic promises of the Millennium Declaration. To achieve sustainable outcomes, country-led development strategies must be backed by adequate financing within the global partnership for development. However, this is only possible if timely evidence is available from policy-relevant and technically-reliable country-led monitoring and evaluation systems. The evidence provided by such systems, owned by developing and transition countries, should inform necessary policies and strategies to ensure progress.

## DevInfo is being used to support country-led monitoring and evaluation systems

DevInfo is a database system which harnesses the power of advanced information technology to compile and disseminate data on human development. In particular, the system has been endorsed by the UN Development Group to assist countries in monitoring achievement of the Millennium Development Goals (MDGs).

DevInfo provides methods to organize, store and display data in a uniform way, to facilitate data sharing at the country level across government departments, UN agencies and development partners. DevInfo has simple and user-friendly features which produce tables, graphs and maps for inclusion in reports, presentations and advocacy materials. The software supports both standard indicators (the MDG indicators) and user-defined indicators. DevInfo is compliant with international statistical standards to support open access and widespread data exchange. DevInfo is distributed royalty-free to all Member States and UN agencies, for deployment on both desktops and the web. The user interface of the system, as well as the contents of the databases supported by the system, include country-specific branding and packaging options. These options have been designed for broad ownership by national authorities.

The vision that DevInfo supports is a day when Member States use common database standards for tracking national human development indicators, containing high-quality data with adequate coverage and depth, to sustain good governance around the agenda of achieving the MDGs and national development goals.

DevInfo is being used as an advocacy platform to engage a broad spectrum of stakeholders in policy choices for human development. Member States and UN agencies around the world are using DevInfo to help support the reform of development planning policies. The system is enabling the UN to work together as "One UN" and to effectively deliver as one UN system based on a common database that leads to a common understanding of how to move forward together, with less duplication of efforts and wasteful delays in progress.

DevInfo is being used as a tool to restructure programming processes based on human rights. The system helps planners address disparities and target the most vulnerable sections of society. An important aspect of the DevInfo database structure is that it provides for monitoring multiple levels of sub-national data. The database structure also provides methods for monitoring subgroups (by sex, location (urban/rural), age-groups, ethnicity, education level, wealth index), and other important factors related to groups at risk and in need.

DevInfo can help design cost effective interventions based on facts, not perceptions. The system helps planners evaluate their options to plan for optimum results with limited resources. DevInfo presents the facts from multiple data sources with extensive metadata. This

assists planners to assess all of the available data related to the current situation, weigh alternatives and plan ahead as effectively as possible.

## **DevInfo. A database system designed to facilitate ownership by national authorities**

### ***National ownership and demand-driven monitoring and evaluation systems***

Progress in human development is being made even in countries where the challenges are the greatest. This progress testifies to the unprecedented degree of commitment by these countries to achieve results through national ownership of the development process. National ownership of data dissemination processes helps to ensure that all stakeholders can make informed decisions about the future course of development policies that affect them as individuals, communities and the nation as a whole.

A survey conducted by UNICEF CEE/CIS Regional Office in 2008 showed that 68% of countries in the region are in various stages of DevInfo implementation. In most of these countries, the National Statistics Office (NSO) is the owner of the database, while in 32% of them the ownership is shared with other agencies or ministries. For example, in Kosovo, the Ministry of Science and Technology is supporting the DevInfo initiative. In Tajikistan, the Ministry of Economic Development and Trade is a national partner, along with the NSO.

The selection of indicators contained in a DevInfo database is demand-driven. This ensures that a national database will sustain its relevance and importance as a useful tool for monitoring national frameworks. The data's relevance, for tracking these frameworks, is critical to the success of the implementation of the database system. Successful DevInfo implementations have identified stakeholders and ensured their participation in governance of the system. The stakeholders have thoroughly examined the legal framework for gathering and use of statistics in the country, and its ramifications for DevInfo. They have leveraged relevant institutional structures and processes of government and partners to strengthen national data dissemination. Considering these issues helps position DevInfo strategically, creating links to relevant activities, such as in the areas of national strategic planning and support to the statistical system in the country. In this way DevInfo is conceived as a component of a more strategic approach to achieve national development goals.



DevInfo is being used by Member States to monitor comprehensive plans for sustainable development, including poverty reduction strategies, health and nutrition plans, environmental plans and education plans. DevInfo is being implemented by complementing existing databases and bridging data dissemination gaps.

Most of the countries in the CEE/CIS region that are implementing DevInfo have not limited the content of the national databases to the monitoring of the MDGs. Albania, Armenia, Bosnia and Herzegovina, Moldova and Serbia expanded its scope to monitor national development strategies, including poverty reduction strategies (PRSPs). Albania and Turkey are using DevInfo to monitor EU-related strategies, including social exclusion. In some cases DevInfo is being used for monitoring sectoral strategies, such as health care reform in Kyrgyzstan and the education strategy in Kosovo.

Picture 1: *ArmeniaInfo*, national adaptation in Armenia, is used to monitor MDGs as well as national development strategies



There are more than 16 national adaptations of DevInfo database technology in the CEE/CIS region. Some of these adaptations have been deployed online: for example, Tajikistan launched *TajikInfo* at [www.tajikinfo.tj](http://www.tajikinfo.tj) and Moldova launched *MoldovaInfo* at [www.devinfo.md](http://www.devinfo.md). Four national databases (Armenia, Azerbaijan, Macedonia and Serbia) are hosted at the global DevInfo website [www.devinfo.info](http://www.devinfo.info). In addition, the websites of the national statistical offices of Serbia (<http://webzrs.statserb.sr.gov.yu/axd/devinfo/indexe.htm>) and Montenegro ([www.monstat.cg.yu/EngProjekti.htm](http://www.monstat.cg.yu/EngProjekti.htm)) allow users to download their databases to function with the desktop version of DevInfo.

Picture 2: *TojikInfo*, local adaptation in Tajikistan, is available on line.



Picture 3: *Kyrgyzstan HealthInfo*, local adaptation in Kyrgyzstan, is used to monitor health reform.



National ownership processes entail several elements. It starts from the signature of a Memorandum of Understanding among stakeholders, to build a common database to monitor national development priorities. It then moves on to: outline roles and responsibilities of all stakeholders; commit financial and human resources; establish a steering committee to govern the content of the database; assign working groups to update the database; decide on the location of the common database; and finally, to end up with the integration of DevInfo database technology into the internal infrastructure of the government. This results in full institutionalization of the system.

An example of full ownership of the DevInfo system by a government is in the case of the Republic of Serbia. The government declared DevInfo as a database tool of particular interest for the Republic of Serbia in 2006. The technology thereby became part of the regular programme of the Statistical Office of the Republic of Serbia (SORS). This led to the formation of a committee on social indicators and analysis. The unit consists of four people, supported by the government, who have undertaken the task of further development and maintenance of the DevInfo database at the national

level. As a result, the national DevInfo database contains a rich set of 395 indicators at national level, which are classified in 12 sectors with 5 multilateral strategies: Millennium Development Goals (MDGs); Poverty Reduction Strategy (PRS); National Plan of Action for Children (NPA); World Fit for Children; and, World Summit for Children. The database also contains data on 91 indicators at local level (for each of 167 municipalities). A specially designed census database has 62 indicators at the settlement level (for each of 4,715 settlements). These databases are strong tools for monitoring and planning at central and local level.

Important initiatives are also taking place in other regions. For example, the Costa Rica government selected a strategic implementing partner, made them responsible for the system, so they took ownership and so, are developing it further, promoting it, and most importantly, sharing the information it contains.

In Egypt, a Memorandum of Understanding was signed among government agencies in charge of data collection, processing, analysis and dissemination. A major advantage is the linkage of DevInfo adaptations to existing decision-making mechanisms and processes in the country. For that purpose, it is helpful for a government body, directly linked to the decision-making process, to manage the system.

Tanzania's *TSED*, for example, is owned by the National Bureau of Statistics in collaboration with more than 20 ministries, departments and agencies in the country. It is embedded in the monitoring system for the National Strategy for Growth and Reduction of Poverty. In order to ensure the relevance of Tanzania's *TSED*, the database includes data for: the MDGs; the country's National Strategy for Growth and Reduction of Poverty; and, other relevant frameworks, such as Ageing and Aged Population; Labor Market Indicators; Maternal and Child Monitoring Indicators; and, Education for All. In addition, the National Bureau of Statistics implements a process for ensuring the quality, accuracy and reliability of the data. These conditions encourage the use of the database to produce reports to monitor the National Strategy for Growth and Reduction of Poverty, and it enables the government and its partners to gauge the progress being made by various interventions. Civil society organizations are using *TSED* in advocacy work related to policy formulation and budgetary processes. Others have also used the database for reporting, proposal writing and presentations.

Cambodia provides a clear illustration of strategic linkages. The Statistical Literacy Project has partnered with the *CAMInfo* initiative to conduct joint nation-wide trainings on *CAMInfo* and statistical literacy, targeting government officials and users of statistical data, including high-level decision makers. This partnership is expected to promote better coordination between the data manager, the National Institute of Statistics, and the planning and decision-making agency, the Ministry of Planning. As a result, better access to quality data and improved statistical literacy are anticipated to contribute to the improvement of the government's capacity to integrate statistical information into policy making. In St. Lucia, *Helen Info* is designed to be used by the government for Evidence-Based Social Policy. The database has been established in partnership between Government, EU, UNDP and UNICEF. Most important has been government ownership and their commitment to maintain and use the database. Following this successful example, DevInfo is now being rolled out throughout the Eastern Caribbean.

### ***National capacity development***

Access to timely and reliable development data plays an important role in helping identify national development issues and, through national capacity development in data dissemination, leads to better information for policy development. Progress is being made in sharpening national monitoring and evaluation systems and this is enhancing the impact of development funding. These efforts are being stepped up to increase awareness of potential problems and to find solutions for extreme disparities and vulnerabilities. Since 2004, more than 20,000 professionals have been trained in the use of DevInfo database technology. These training sessions have focused on best practices in establishing a common database on human development and on how to put the data to use for decision making. The training has targeted a broad audience of planners, politicians, policy analysts, researchers, teachers, youth and statisticians. It has been organized at global, regional, national and local levels. The strategy has been to create teams of master trainers who can assist others to become both trainers and database administrators.

National capacity development is also provided through technical missions and activities to assist national partners and UN agencies in setting up and using DevInfo database technology. In 2007, there were 298 technical support activities carried out. This has resulted in more than 120 countries using DevInfo as the database platform to develop their own national socio-economic databases.

Capacity development activities in Central and Eastern Europe and the Commonwealth of Independent States (CEE/CIS) started with a series of DevInfo roll-out training carried out by the UNICEF CEE/CIS Regional Office. The scope of this training varied from orientation and use of the software to advanced database administration and development of local adaptations of the database technology to meet country-specific requirements. There was also a session devoted to Training of Trainers in the user and data administration modules of DevInfo.

Since 2006 regional training has been implemented in partnership with the United Nations Economic Commission for Europe (UNECE) and UNDP Bratislava Regional Center. The training introduced DevInfo v5.0, a new version with the capability of disseminating data online. The DevInfo regional training brought together national partners and UN staff members already working together on monitoring national development priorities. These regional capacity building activities have been supplemented by the UN Development Group Office (UNDGO, now UNDOCO) which facilitated training in priority countries and included the roll-out of the UN Development Assistance Framework (UNDAF). These training activities were organized through the countries' UN Resident Coordinators.

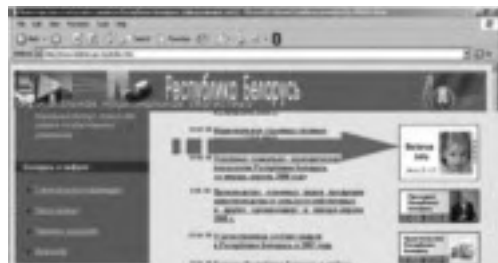
Promoted by these regional activities, much in-country training has been carried out. According to an e-mail survey carried out by the UNICEF CEE/CIS Regional Office in February 2008, more than 1000 people in CEE/CIS have been trained in DevInfo. This provides a critical mass of technical capacity to convey knowledge about the system and to carry out national and sub-national training.

In-country training is vital to the implementation of DevInfo database technology. This training, organized on behalf of national authorities, is integrated into a broad framework for monitoring national development priorities. Training focuses on the demand for data to monitor local circumstances.

An example of national capacity building is the step-by-step introduction of DevInfo in the Republic of Belarus. It started with a needs assessment in 2005, followed by participation in the DevInfo 5.0 regional roll-out training in Geneva (2006). The regional roll-out training was followed by a country request to carry out a session on DevInfo database administration in Belarus. This covered an overview for a wider international and national community and hands-on training for Ministry of Statistics and Analysis staff members. In 2006, database administration training was attended by 22

participants. This was facilitated in Russian by the UNICEF Regional Office, in collaboration with the UNDP and UNICEF country offices, and with the technical and logistical support of the Ministry of Statistics and Analysis. As a result of the training, the Ministry finalized a national adaptation of DevInfo for Belarus in 2007. The current version of *BelarusInfo* contains 126 indicators, focuses on national MDGs and provides access to socio-economic indicators related to human development in the country.

Picture 4: *BelarusInfo* is accessible at the website of the Ministry of Statistics and Analysis of the Republic of Belarus



Information on *BelarusInfo* can be obtained at [www.belstat.gov.by](http://www.belstat.gov.by). The database is currently available in Russian. The Ministry of Statistics and Analysis, in collaboration with UNDP and UNICEF, is plans to update, translate and further disseminate *BelarusInfo*, to insure wide access and usage of the database for informed decision making on national and the sub-national levels. Sub-national level training is also being planned.

### ***Monitoring UN contribution to national development strategies and priorities***

The United Nations Development Assistance Framework (UNDAF) is the strategic programme framework for the national development strategies supported by the UN Country Team. It describes UN contribution to the priorities in the national development framework. The outcomes of the framework show where the UN Country Team can bring its unique comparative advantages to bear in advocacy, capacity development, policy advice and programming for the achievement of related national priorities. A successful UNDAF is dependent on a strong, relevant national data dissemination system.

In India, the features of *DevInfo India* are being implemented to generate information on the overall situation with respect to sustainable development. The monitoring framework is inclusive of indicators to measure UNDAF outcomes/outputs, information on trends/mechanism for coordination, tracking of national development over time, progress of joint-sector programmes and responses to humanitarian emergencies. In Lesotho, *MalutiInfo* helps make information easily accessible to policy-makers, development practitioners and others, thus allowing them to monitor and evaluate the performance of identified indicators related to the UNDAF, PRS and MDGs. To increase the usefulness of the database, the country has created report templates to generate regular progress reports on thematic development agendas such as those related to the UNDAF; UN Common Country Assessment; National Human Development Reports; and, the Situational Analysis of Women and Children. Similarly, Malawi's *MASEDA* contains indicators for monitoring the country's development strategies, MDGs, and the UNDAF monitoring and evaluation (M&E) matrix, supplemented by indicators from other relevant areas such as governance. In Cambodia, *CAMIInfo* was adapted to include not only the indicators specific to monitoring the UNDAF, but additional indicators in the areas of governance and human rights, in order to capture more qualitative information and results at the output/outcome level.

### ***Local monitoring and evaluation systems to strengthen decentralization***

Successful national development strategies are built on sound economic and technical information which are used to design programmes to overcome key development challenges. These strategies are aimed to reduce child and maternal mortality, extreme poverty, lack of basic sanitation, unemployment and increasing inequalities. To be effective, national development strategies must be universal while targeting the most vulnerable and marginalized to reduce disparities. Policymakers must know where disparities exist within their own countries in order to develop relevant solutions which benefit the poor. The poor are often those living in rural areas or urban slums, children of mothers with no formal education, and living in the poorest households. National monitoring and evaluation systems focusing on disaggregated data, as well as decentralized systems, are fundamental to provide the information needed for policy makers to design and implement such developing strategies.

In Albania, UNDP (in partnership with UNICEF and UNFPA), supported local authorities, in all 12 regions of Albania, in developing Regional Development Plans. The decentralized monitoring and evaluation system is being supported by DevInfo. In Serbia, in compliance with the National Plan of Action for Children, 16 municipalities initiated Local Plans of Action for Children (LPA). These are strategic documents to define and guide optimal child development in local settings. The municipalities have been introduced to DevInfo to monitor progress, assess the local situation and inform decision making. Similarly, municipal databases are being developed in Montenegro. In Bosnia and Herzegovina, ten municipalities are working on the adaptation of DevInfo to strengthen child rights monitoring. In some municipalities, DevInfo is also used for monitoring the child protection systems reform. Data from municipalities is being sent to the Department of the Economic Development at central level where a consolidated dataset is used for national level planning and fund allocation. In the Russian Federation, the municipality of Moscow is exploring the opportunity of using DevInfo to monitor the Child Friendly Cities Initiative.

## **DevInfo is being used to monitor regional development challenges**

DevInfo is being used at transnational level to highlight and monitor specific development challenges common to a group of countries or regions. For example, the UNICEF CEE/CIS Regional Office developed three adaptations: *MONEEInfo*, *MICS Info* and *Regional MGDInfo*.

*MONEEInfo* – available in online at [www.moneeinfo.org](http://www.moneeinfo.org) – consists of 128 indicators related to the MDGs and beyond. *MONEEInfo*, based on the UNICEF IRC TranMonee database, allows monitoring of the situation of women and children in 27 countries of the region using time series from 1989 to the most recent year for which data are available. It is available in Russian and English. *MONEEInfo* provides a rich resource to access and analyze child protection indicators related to the institutionalization of children, living arrangements and juvenile justice, among other related issues.

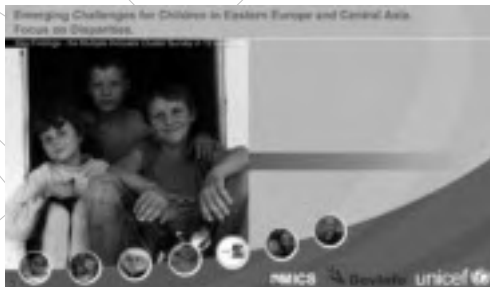


Picture 5: *MONEE Info*, a regional adaptation developed by UNICEF CEE/CIS Regional Office, is based on TransMONEE data



*MICSInfo* (accessible at [www.micsinfo.org](http://www.micsinfo.org)), presents the findings for the third round of Multiple Indicator Cluster Surveys carried out in 12 countries of the CEE/CIS region. This DevInfo adaptation consists of a DevInfo gallery provides access to the charts with the key findings; the downloadable tables; the report “*Emerging challenges for children in Eastern Europe and Central Asia – Focus on disparities*”; and, provides full access to data on 59 indicators, including new indicators on child protection and early childhood development. Data are disaggregated by age, gender, family size, children living arrangement, residence, mother’s education, wealth index and ethnicity/language/religion.

Picture 6: *MICSInfo*, a regional adaptation developed by UNICEF CEE/CIS Regional Office, presents MICS3 data



The *Regional MDGInfo* database – accessible at [www.regional-mdginfo.org](http://www.regional-mdginfo.org) – has been developed through partnership of UNICEF, UNDP and UNECE in an effort to strengthen national capacities in MDG literacy and monitoring. The database is used in advocacy for improvements in data quality and comparability. There are 78 indicators stratified by different background variables in the database. The gallery provides easy access to presentations of the key find-

ings related to progress towards the MDGs. *Regional MDGInfo* contains indicators from both national and international sources, as well as regionally-specific indicators, to maximize the relevance of MDG monitoring to the national context and to promote evidence-based advocacy for policy making.

Picture 7: Regional MDGInfo was developed by UNICEF, UNECE and UNDP



## Data disseminated through DevInfo contributed to achieving results for children

Most of the countries in the CEE/CIS region that are using DevInfo report that the system is being used for preparing progress reports on MDGs and national development strategies. Serbia and Moldova reported that DevInfo was able to trigger important policy changes, including in public budgets, both at national and decentralized level.

According to Salah (2008), in Moldova, the DevInfo database of the Ministry of Economy and Trade provides central public authorities with relevant and internationally comparable statistical data on a regular basis. By using the same technology and the same lists of indicators in building two integrated national databases – Economic Growth and Poverty Reduction Strategy database (EGPRSP), and MDG database – the team avoided duplication in collecting statistics and increased the reliability of reporting. They also avoided the complexity which traditionally occurs in maintaining statistical data systems. With the objective of improving national capacity in decision-making, the Ministry of Economy and Trade developed two different types of comprehensive, analytical reports which are also DevInfo based. One, the Annual Evaluation Report on the Implementation of the Economic Growth and Poverty Reduction Strategy Paper, helped social sector ministries to discuss budgetary questions with the Ministry

of Finance. As a result, investments in social sectors were raised by 21 per cent in 2006. The other, the 2005 Poverty and Policy Impact Report, provided an overview of national development and included detailed analyses on child poverty and on poverty in rural areas.

These reports did not replace economic evaluations and public expenditures reviews. They did however provide useful information for decision-making since they contained analyses which indicated those elements which influenced programme results, and how the programme elements interacted among themselves. The reports were produced through an inclusive and nationally owned process where staff from MoET interacted with key decision-makers in line ministries. Because they provided objective analyses of local realities, they were also used by external donors. MoET organized an annual event which was a major opportunity for an evidence-based and participatory reflection on Moldova's performance in the economic and social sectors, and for a comparison with other countries. The reports were used for strategic planning including by teams developing the National Development Plan (NDP) 2008–2011. DevInfo played a role in facilitating a common understanding among the government, civil society organizations (CSOs) and development partners. Data analyses and maps were used as platforms for the national dialogue on poverty reduction. As information was easily accessible, DevInfo was used to produce a bulletin on EGPRSP implementation which was published in Moldovan newspapers and posted on government websites. This bulletin led to increased CSO participation and involvement in EGPRSP implementation. The materials developed by MoET for monitoring the *Poverty Reduction Strategy* helped a coalition of 14 non-government organizations (NGOs) develop the *State of the Nation Report* which presented civil society's view of development in Moldova. The main purpose of the Report was to play a role in decision-making and, in particular, to influence the content of the NDP for 2008–2011.

At the decentralized level, the municipality of Pirot in Serbia (Vasic, Petrovic and Jancovic, 2008) used DevInfo for reviewing the municipal budget allocation in favor of children. As a result, investment for children was increased seven-fold in just two years starting in 2005. In addition, an increasing demand from the local population for better quality of child social services prompted local authorities to provide additional funds. Firstly, additional funds were invested to equip the antenatal service. Secondly, there was increased funding of the Social Welfare Centre, schools and NGOs. Additionally, a new pre-school was built which tripled access to early childhood

education, raising it to 90% in the municipality. In the same municipality, DevInfo enabled local government to identify that none of the Roma children were attending pre-school facilities and that most of the children in the specialized institutions for children with disabilities were Roma. As a result, 50 children from Roma settlements were enrolled into pre-school (rather than in specialist institutions), and in one school year the proportion of Roma children in specialized institutions was reduced by 50%.

In Bosnia and Herzegovina, data disseminated through DevInfo are producing policy changes in education. Previously municipal authorities thought enrollment to primary school was 100 per cent. Now, thanks to data disseminated through DevInfo, local authorities realized that the situation is different for marginalized children. DevInfo also helped local municipalities to have a better insight in the area of social protection services, including for vulnerable and excluded groups, as well as on municipal budget allocation for children.

## Conclusions

The DevInfo database initiative is proving that progress in human development can be accelerated through nationally-owned systems to strengthen data dissemination. The progress being made in use of data for decision-making bears witness to the unparalleled degree of advancement that can be achieved through ready access to relevant development data.

DevInfo is being used by the United Nations to strengthen its strategic national programme frameworks to deliver as One UN based on new approaches to create a common database on human development indicators supported by a strong data dissemination system.

National ownership of such data dissemination system is vital to the future course of human development where all stakeholders are able to be actively involved in evidence based policy decision making processes.

## References

Djokovic-Papic, Dragana, Oliver Petrovic and Vladica Jankovic. (2008). *Using DevInfo to support Governments in monitoring National Development Strategies. The case of the Republic of Serbia*. In: Bridging the gap. The role of monitoring and evaluation in evidence-based policy making. UNICEF CEE/CIS Regional Office, Switzerland.

Pron, Nicolas. (2008). *The strategic intent of data dissemination. The case of DevInfo*. In: Bridging the gap. The role of monitoring and evaluation in evidence-based policy making. UNICEF CEE/CIS Regional Office, Switzerland.

Salah, Mohamed Azzedine. (2008). *Using DevInfo as a strategic tool for decision making. Achievements and lessons learned in Moldova*. In: Bridging the gap. The role of monitoring and evaluation in evidence-based policy making. UNICEF CEE/CIS Regional Office, Switzerland.

Segone, Marco. (2008). *Evidence-based policy making and the role of monitoring and evaluation within the new aid environment*. In: Bridging the gap. The role of monitoring and evaluation in evidence-based policy making. UNICEF CEE/CIS Regional Office, Switzerland.

Vasic, Vladan, Oliver Petrovic and Vladica Jankovic. (2008). *Using DevInfo as a strategic tool to facilitate local communities' empowerment. The case of the Municipality of Piro*. In: Bridging the gap. The role of monitoring and evaluation in evidence-based policy making. UNICEF CEE/CIS Regional Office, Switzerland.

United Nations. (2007). *The Millennium Development Goals Report. 2007*.

# MAKING DATA MEANINGFUL. WRITING STORIES ABOUT NUMBERS.<sup>1</sup>

**Colleen Blessing**, *United States Department of Energy*  
**Vicki Crompton**, *Statistics Canada*  
**Dag Ellingsen**, *Statistics Norway*  
**Patricia Fearnley**, *Office for National Statistics, United Kingdom*  
**John Flanders**, *Statistics Canada*  
**John Kavalionas**, *United States Census Bureau*  
**David Marder**, *Office for National Statistics, United Kingdom*  
**Steve Matheson**, *Australian Bureau of Statistics*  
**Kenneth Meyer**, *United States Census Bureau*  
**Hege Pedersen**, *Statistics Norway*  
**Sebastian van den Elshout**, *Statistics Netherlands*  
**Don Weijers**, *Statistics Netherlands*  
**Marianne Zawitz**, *United States Bureau of Justice Statistics*

---

*Making Data Meaningful. A guide to writing stories about numbers* was prepared within the framework of the United Nations Economic Commission for Europe (UNECE) Work Session on Statistical Dissemination and Communication, under the programme of work of the Conference of European Statisticians.

The guide is intended as a practical tool to help managers, statisticians and media relations officers use text, tables, graphics and other information to bring statistics to life using effective writing techniques. It contains suggestions, guidelines and examples – but not golden rules. This publication recognizes that there are many practical and cultural differences among statistical offices, and that approaches vary from country to country.

## What is a statistical story?

On their own, statistics are just numbers. They are everywhere in our life. Numbers appear in sports stories, reports on the economy, stock market updates, to name only a handful. To mean anything, their value to the person in the street must be brought to life.

---

<sup>1</sup> *Making Data Meaningful: A guide to writing stories about numbers* was originally published by the United Nations Economic Commission for Europe (UNECE). Reprinted with the permission of UNECE.

A statistical story is one that doesn't just recite data in words. It tells a story about the data. Readers tend to recall ideas more easily than they do data. A statistical story conveys a message that tells readers what happened, who did it, when and where it happened, and hopefully, why and how it happened. A statistical story can:

- provide general awareness/perspective/context; and
- inform debate on specific issues.

In journalistic terms, the number alone is not the story. A statistical story shows readers the significance, importance and relevance of the most current information. In other words, it answers the question: *Why should my audience want to read about this?*

Finally, a statistical story should contain material that is newsworthy. Ask yourself: Is the information sufficiently important and novel to attract coverage in the news media? The media may choose a different focus. But they have many other factors to consider when choosing a story line.

Statistical story-telling is about:

- catching the reader's attention with a headline or image;
- providing the story behind the numbers in an easily understood, interesting and entertaining fashion, and;
- encouraging journalists and others to consider how statistics might add impact to just about every story they have to tell.

## Why tell a story?

A statistical agency should want to tell a story about its data for at least two reasons. First, the mandate of most agencies is to inform the general public about the population, society, economy and culture of the nation. This information will guide citizens in doing their jobs, raising their families, making purchases and in making many other decisions. Secondly, an agency should want to demonstrate the relevance of its data to government and the public. In such a way, it can anticipate greater public support for its programmes, as well as improved respondent relations and greater visibility of its products.

Most agencies rely mainly on two means of communicating information on the economic and social conditions of a country and its citizens: the Internet and the media. The Internet has become an

important tool for making access easier to the agency's information. More and more members of the public access an agency's data directly on its website. Still, most citizens get their statistical information from the media, and, in fact, the media remain the primary channel of communication between statistical offices and the general public. An effective way for a statistical office to communicate through both means is to tell a statistical story that is written as clearly, concisely and simply as possible. The goal for the Internet is to better inform the public through direct access. When writing for the media, the aim is to obtain positive, accurate and informative coverage. Statistics can tell people something about the world they live in. But not everyone is adept at understanding statistics by themselves. Consequently, statistical stories can, and must, provide a helping hand.

Last, but certainly not least, the availability of statistics in the first place depends on the willing cooperation of survey respondents. Statistical agencies cannot just rely on their legal authority to ensure a suitable response rate. The availability of statistics also depends on the extent to which survey respondents understand that data serve an important purpose by providing a mirror on the world in which we live. The more a statistical agency can show the relevance of its data, the more respondents will be encouraged to provide the data.

## **Considerations**

Statistical agencies must take into account a number of key elements in publishing statistical stories.

First, the public must feel that it can rely on its national statistical office, and the information it publishes. Statistical stories and the data they contain must be informative and initiate discussion, but never themselves be open to discussion. In other words, the information must be accurate and the agency's integrity should never come into question. Statistical agencies should always be independent and unbiased in everything they publish. Stories must be based on high-quality data which are suitable to describe the issues they address. Changes in statistical values over time, for example, should be discussed only if they are determined by statisticians to be statistically significant.

Agencies should always guarantee the confidentiality of data on individual persons or businesses. Indeed, statistical stories may not iden-



tify, or in any way reveal, data on individuals or businesses. In their statistical storytelling, agencies must take into account the position and feelings of certain vulnerable groups in society. Information on these groups should be made available, but the goal should always be to inform the public. Agencies should never seek publicity for themselves at the expense of these particular target groups.

The authors of this guide suggest that statistical agencies should, for the benefit of the citizens they serve, formulate a policy that explains how their practices protect the privacy and confidentiality of personal information. This policy should be given a prominent position on the agency's website.

## **How to write a statistical story**

### ***Do you have a story?***

First and foremost, you need a story to tell. You should think in terms of issues or themes, rather than a description of data. Specifically, you need to find meaning in the statistics. A technical report is not a story, nor is there a story in conducting a survey. A story tells the reader briefly what you found and why it is important to the reader. Focus on how the findings affect people. If readers are able to relate the information to important events in their life, your article becomes a lot more interesting.

Statistical offices have an obligation to make the data they collect useful to the public. Stories get people interested in statistical information and help them to understand what the information means in their lives. After they read good statistical stories, people should feel wiser and informed, not confused.

Possible topics/themes for stories:

- current interest (policy agenda, media coverage, etc.);
- reference to everyday life (food prices, health, etc.);
- reference to a particular group (teens, women, the elderly, etc.);
- personal experiences (transportation, education, etc.);
- holidays (Independence Day, etc.);
- current events (statistics on a topic frequently in the news);
- calendar themes (spring, summer, etc.);
- new findings;

- a regular series (“This is the way we live now”, “Spotlight on xxxx”, etc.).

### ***Write like a journalist. The “inverted pyramid”***

How can statisticians communicate like journalists? By writing their stories the way journalists do. The bonus is that the media are more likely to use the information.

Journalists use the “inverted pyramid” style. Simply, you write about your conclusions at the top of the news story, and follow with secondary points in order of decreasing importance throughout the text. Think of a typical analytical article as a right-side-up pyramid. In the opening section, you introduce the thesis you want to prove. In following sections, you introduce the dataset, you do your analysis and you wrap things up with a set of conclusions. Journalists invert this style. They want the main findings from those conclusions right up top in your news story. They don’t want to have to dig for the story.

You build on your story line throughout the rest of the text. If the text is long, use subheadings to strengthen the organization and break it into manageable, meaningful sections. Use a verb in subheadings, such as: “Gender gap narrows slightly.”

### ***The lead. Your first paragraph***

The first paragraph, or lead, is the most important element of the story. The lead not only has to grab the reader’s attention and draw him or her into the story, but it also has to capture the general message of the data. The lead is not an introduction to the story. On the contrary, it should tell a story about the data. It summarizes the story line concisely, clearly and simply. It should contain few numbers. In fact, try writing the first sentence of the lead using no figures at all.

Don’t try to summarize your whole report. Rather, provide the most important and interesting facts. And don’t pack it with assumptions, explanations of methodology or information on how you collected the data.

The lead paragraph should also place your findings in context, which makes them more interesting. Research shows that it is easier to remember a news report if it establishes relevance, or attempts to explain a particular finding. Exercise caution, though. It is not a good

idea to speculate, especially if your statistical office cannot empirically establish causality, or does not produce projections.

Give enough information so the reader can decide whether to continue reading. But keep it tight. Some authors suggest five lines or fewer – not five sentences – for the opening paragraph.

**Poor:** A new study probes the relationship between parental education and income and participation in post-secondary education from 1993 to 2001.

**Good:** Despite mounting financial challenges during the 1990s, young people from moderate and low-income families were no less likely to attend university in 2001 than they were in 1993, according to a new study.

Finally: there is no contradiction between getting attention and being accurate.

Remember:

- focus on one or two findings;
- write in everyday language (the “popular science” level);
- create images for your readers;
- focus on the things you want readers to remember;
- choose the points you think are newsworthy and timely.

### ***Good writing techniques***

Write clearly and simply, using language and a style that the layperson can understand. Pretend you are explaining your findings to a friend or relative who is unfamiliar with the subject or statistics in general. Your readers may not be expert users who often go straight to the data tables. Terms meaningful to an economist may be foreign to a layperson, so avoid jargon. Use everyday language as much as possible. If you have to use difficult terms or acronyms, you should explain them the first time they are used.

Remember: on the Internet, people want the story quickly. Write for the busy, time-sensitive reader. Avoid long, complex sentences. Keep them short and to the point. Paragraphs should contain no more than three sentences.

Paragraphs should start with a theme sentence that contains no numbers.

Example: Norway’s population had a higher growth last year than the year before. The increase amounted to 33,000 people, or a growth rate of 0.7%.

Large numbers are difficult to grasp. Use the words millions, billions or trillions. Instead of 3,657,218, write “about 3.7 million.” You can also make data simpler and more comprehensible by using rates, such as per capita or per square mile. Some suggestions follow.

<i>Use</i>	<i>Avoid</i>
<ul style="list-style-type: none"> <li>• Language that people understand;</li> <li>• Short sentences, short paragraphs;</li> <li>• One main idea per paragraph;</li> <li>• Subheadings to guide the reader’s eye;</li> <li>• Simple language: “Get,” not “acquire.” “About,” not “approximately.” “Same,” not “identical”;</li> <li>• Bulleted lists for easy scanning;</li> <li>• A good editor. Go beyond Spell-Check; ask a colleague to read your article;</li> <li>• Active voice. “We found that...” Not: “It was found that....”;</li> <li>• Numbers in a consistent fashion: For example, choose 20 or twenty, and stick with your choice;</li> <li>• Rounded numbers (both long decimals and big numbers);</li> <li>• Embedded quotes (these are sentences that generally explain “how” or “why”, and which journalists like to use verbatim in their news stories in quotes);</li> <li>• URLs, or electronic links, to provide your reader with a full report containing further information.</li> </ul>	<ul style="list-style-type: none"> <li>✗ “Elevator statistics”: This went up, this went down, this went up;</li> <li>✗ Jargon and technical terms;</li> <li>✗ Acronyms;</li> <li>✗ All capital letters and all italics: Mixed upper and lower case is easier to read;</li> <li>✗ “Table reading”, that is, describing every cell of a complex table in your text.</li> </ul>

Not Good: From January to August, the total square metres of utility floor space building starts rose by 20.5% from the January to August period last year.

Better: In the first eight months of 2004, the amount of utility floor space started was about 20% higher than in the same period of 2003.

### ***Headlines. Make them compelling***

If your agency's particular style calls for a headline on top of a statistical story, here are some suggestions to keep in mind.

Readers are most likely to read the headline before deciding to read the full story. Therefore, it should capture their attention. The headline should be short and make people want to read on. It should say something about the findings presented in the article, not just the theme.

Write the headline after you have written your story. Headlines are so important that most newspapers employ copy editors who craft the headlines for every story. Because the information is likely to be new to them, these editors can focus more readily on the most interesting aspects of the story.

In the same vein, statistical agencies might consider a similar arrangement. The individual who writes the headline could be different than the story's author.

Headlines should:

- be informative, appealing, magnetic, interesting and newsy, and incorporate:
  - the highest since, the lowest since...;
  - something new;
  - the first time, a record, a continuing trend;
- make you want to read the story, not scare you off;
- summarize the most important finding;
- be no longer than one line of type;
- not try to tell everything;
- contain few numbers, if any at all;
- have a verb or implied verb.

- Not Good: New report released today (the report is not the news)  
Energy conservation measures widespread (too vague)  
Prices up in domestic and import markets (what prices?)
- Good: Gasoline prices hit 10-year low  
Crime down for third year in a row  
July oil prices levelled off in August

### ***Tips for writing for the Internet***

The principles of good writing also apply to writing for the Internet, but keep in mind some additional suggestions.

People scan material on the Internet. They are usually in a hurry. Grabbing their attention and making the story easy to read are very important. You also have different space limitations on the Internet than on paper. Stories that make the reader scroll through too many pages are not effective. Avoid making the reader scroll horizontally.

Format the page so the story can be printed properly, without text being cut off by margin settings. A common solution is to include a link to a 'print friendly version', usually another page with navigation menus and banners removed.

Write your text so the reader can get your point without having to force themselves to concentrate. Use structural features such as bulleted lists, introductory summaries and clear titles that can stand alone.

Don't use ALL CAPITAL LETTERS on the Internet. It looks like you're shouting. Underline only words that are electronic links. Use boldface rather than underlining for emphasis. Avoid italic typefaces because they are much harder to read.

Make sure your story is printed on a contrasting background colour: either light lettering on a dark background or the reverse. High contrast improves readability on the Internet. Make sure items are clearly dated so readers can determine if the story is current.

### ***Graphs***

A picture is indeed worth a thousand words, or a thousand data points. Graphs (or charts) can be extremely effective in expressing key results, or illustrating a presentation.

An effective graph has a clear, visual message, with an analytical heading. If a graph tries to do too much, it becomes a puzzle that requires too much work to decipher. In the worst case, it becomes just plain misleading. Go the extra mile for your audience so that they can easily understand your point.

Good statistical graphics:

- show the big picture by presenting many data points;
- are “paragraphs” of data that convey one finding or a single concept;
- highlight the data by avoiding extra information and distractions, sometimes called “non-data ink” and “chart-junk”;
- present logical visual patterns.

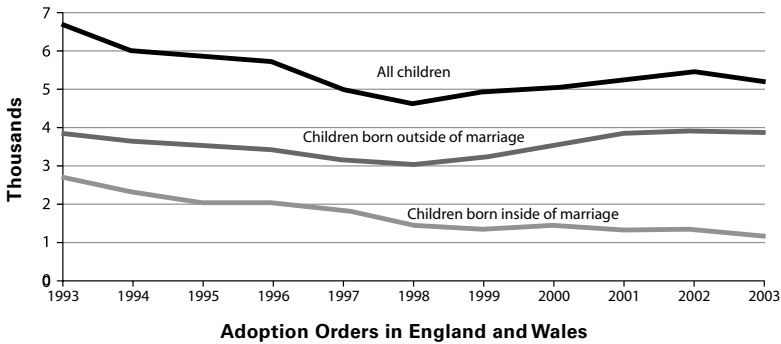
When creating graphics, let the data determine the type of graph. For example, use a line graph for data over time, or a bar graph for categorical data. To ensure you are not loading too many things into a graph, write a topic sentence for the graph.

Achieve clarity in your graphics by:

- using solids rather than patterns for line styles and fills;
- avoiding data point markers on line graphs;
- using data values on a graph only if they don’t interfere with the reader’s ability to see the big picture;
- starting the Y axis scale at zero;
- using only one unit of measurement per graphic;
- using two-dimensional designs for two-dimensional data;
- making all text on the graph easy to understand;
  - not using abbreviations;
  - avoiding acronyms;
  - writing labels from left to right;
  - using proper grammar;
  - avoiding legends except on maps.

For example:

### Adoptions fall by 2.4% in 2003<sup>2</sup>



### Tables

Good tables complement text. They should present numbers in a concise, well-organized fashion to support the analysis. Tables help minimize numbers in the statistical story. They also eliminate the need to discuss insignificant variables that are not essential to the story line.

Make it easy for readers to find and understand numbers in your table. Standard presentation tables are generally small. One decimal place will be adequate for most data. In specific cases, however, two or more decimal places may be required to illustrate subtle differences in a distribution.

Presentation tables rank data by order or other hierarchies to make the numbers easily digestible. They also show the figures that are highest and the lowest, as well as other outliers. Save large complex tables for supporting material. Always right-justify the numbers to emphasize their architecture. The guidelines listed for graphics above, such as highlighting data by avoiding “non-data ink”, also apply to the presentation of tables. While graphics should be accompanied by an analytical heading, titles are preferred for tables. They should be short and describe the table’s precise topic or message.

2 Graph from United Kingdom Office of National Statistics. Available online at <http://www.statistics.gov.uk/cci/nugget.asp?ID=592> [accessed 28 September 2005].



For example:

### ***Race of Juvenile Offenders<sup>3</sup>***

<b><i>Race of juvenile offender(s)</i></b>	<b><i>Average annual percent of violent crimes committed by juvenile(s)</i></b>
<b>Total</b>	<b>100.0%</b>
White	59.1
Black	25.2
Other	11.4
More than 1 racial group	2.6
Unknown	1.7

### ***Maps***

Maps can be used to illustrate differences or similarities across geographical areas. Local or regional patterns, which may be hidden within tables or charts, are often made clear by using a well designed map.

Maps are a rapidly expanding area of data presentation, with methods of geographic analysis and presentation becoming more accessible and easier to use. The cost of Geographic Information Systems (GIS), or software capable of mapping statistics, has decreased rapidly in the last ten years. Mapping that was once expensive, or required specialist hardware, is now within reach of most organizations. GIS analysis and presentation are now taught in schools and universities.

Producing statistical maps can be a simple process. The most common type of statistical map is the choropleth map, where different shades of a colour are used to show contrast between regions (usually a darker colour means a larger statistical value). This type of map is best used for ratio data (e.g. population density), where the denominator is usually area (e.g. square kilometers) or population. 'Count' data which has no denominator (e.g. total number of sheep

3 Table from Juvenile Victimization and Offending, 1993-2003, Bureau of Justice Statistics, Special Report, August 2005, NCJ 209468 (page 8). Available online at <http://www.ojp.usdoj.gov/bjs/pub/pdf/jvo03.pdf> [accessed 28 September]

in each region), are best illustrated using proportional or graduated symbol maps. With proportional symbol maps, the size of a symbol, such as a circle, increases in proportion to the value of the statistic. All mapping software should be capable of producing these two map types. Other types of map are possible but are best retained for specialist audiences.

When designing a map, always think about the audience and try to make it quick and easy for them to understand. If there is a natural association between a colour and a topic (e.g. blue for cold temperatures) then it would be sensible to use that colour for the legend. When choosing your legend classes, do not use complex methods unless your audience will understand them. Choosing classes of equal size, or classes containing similar numbers of events, are the most common methods. When choosing how many coloured classes to use, less is often more. Fewer classes emphasize similarity between areas and more classes emphasize the differences.

It should be possible for any statistical map to be read by a user without reference to other information and knowledge. Maps should always have a title and a legend that adequately explain the statistical units, the date that the statistical information was collected or produced and the geographic area type used. The source of statistical data should also be clearly stated. Footnotes may be used to clarify this information where needed and help to simplify titles.

### Average Annual Rainfall 1961 – 1990, Europe<sup>4</sup>



4 Graph from United Nations Economic Commission for Europe. Available online at <http://www.unece.org/stats/trends2005/environment.htm> [accessed 30 September 2005].

## ***How to encourage good writing***

Each statistical agency may have its own ideas on ways to reward quality writing. But here are some general suggestions.

- set goals, such as a number of stories to be written each year.
- reward good writers for the best headline, most contributions, etc.
- make writing an expected part of the job rather than a sideline.
- explore techniques for building enthusiasm for writing.
- show staff the results of their writing: Post newspaper or magazine coverage initiated by their stories on an office bulletin board.
- provide training.

## **Writing about data. Make the numbers “stick”**

Numbers don’t “talk”. But they should communicate a message, effectively and clearly. How well they do this depends a lot on how well authors use numbers in their text.

In a sense, journalists and statisticians are from two cultures. They tend not to talk the same language. Journalists communicate with words; statisticians communicate with numbers. Journalists are often uncomfortable when it comes to numbers. Many are unable even to calculate a percentage increase. So here are some suggestions for making the data “stick:”

Don’t peel the onion. Get to the point:

Poor: The largest contributor to the monthly increase in the CPI was a 0.5% rise in the transportation index.

Better: Higher auto insurance premiums and air fares helped push up consumer prices this month.

Avoid proportions in brackets:

Poor: Working seniors were also somewhat more likely than younger people to report unpaid family work in 2004 (12% versus 4%).

Better: About 12% of working seniors reported unpaid family work in 2004 compared with 4% for younger people.

Watch percentage changes vs. proportions: A percentage change and a percentage point change are two different things. When you subtract numbers expressed as proportions, the result is a percentage point difference, not a percentage change.

Wrong: The proportion of seniors who were in the labour force rose 5% from 15% in 2003 to 20% in 2004.

Right: The proportion of seniors who were in the labour force rose five percentage points from 15% in 2003 to 20% in 2004.

Avoid changing denominators:

Confusing: Two out of every five Canadians reported that they provided care for a senior in 2001, compared with one in seven in 1996, according to the census.

Clearer: About 40% of Canadians reported that they provided care for a senior in 2001, up from 14% in 1996, according to the census.

Reduce big numbers to understandable levels:

Cumbersome: Of the \$246.8 billion in retail spending last year consumers spent \$86.4 billion on cars and parts, and \$59.3 billion on food and beverages.

Easy to grasp: Of every \$100 spent in retail stores last year, consumers spent \$31 on cars and parts, compared with only \$23 on food and beverages.

### *What's wrong with this article?*

A NEW REPORT RELEASED TODAY SAYS THAT THE PRICES OF MANY PETROLEUM PRODUCTS WILL BE HIGHER IN THE FUTURE

The tight global markets and elevated crude oil prices are expected to result in higher **prices for petroleum products**. The cost of imported crude oil to refineries this winter is projected to average 98.3 c/g (about \$40 per bbl) compared to 70.1 c/g last year. During the winter, WTI prices are expected to

decline from their current record levels but remain in the \$40 per bbl range, but despite above-average natural gas stocks, average winter natural gas prices, both at the wellhead and retail levels, are expected to be above those of last winter, particularly during the fourth quarter of 2004, in response to the hurricane-induced production losses in the Gulf of Mexico during September.

Increases in heating fuel prices are likely to generate higher expenditures even in

regions where demand for fuel is expected to fall. Average residential natural gas prices this winter are expected to be 10 percent higher year-over-year and household expenditures are expected to be 15 percent higher.

Therefore, residential space-heating expenditures are projected to increase for all fuel types compared to year-ago levels.

Demand is expected to be up by 1.637 percent. This increase reflects greater heating degree days in key regions with larger concentrations of gas-heated homes and continued demand increases in the commercial and electric power sectors. Due to the availability of primary inventories, many petroleum products are expected to be reasonably well

protected against the impact of demand surges under most circumstances. As of October 1, working natural gas inventories were estimated to be 3.6tcf, up 2 percent from three years ago, 3 percent from two years ago and 1 percent from last year.

Other interesting findings from this report are that the spot price for crude oil continues to fluctuate. Prices continue to remain high even though OPEC crude oil production reached its highest levels in September since OPEC quotas were established in 1982. Overall inventories are expected to be in the normal range, petroleum demand growth is projected to slow, and natural gas prices will increase.

- Headline is too long and doesn't make a clear point.
- All-cap headline looks like the author is shouting.
- Don't underline words unless they are an electronic link.
- Lead paragraph is background.
- Report title and release date aren't stated.
- Jargon: Readers might not know that gasoline and heating oil are petroleum products.
- Spell out units: c/g is cents per gallon; bbl is barrel.
- Acronyms: OPEC is the Organization of Petroleum Exporting Countries.
- First paragraph is too long: Too much detail, too many numbers.
- Sentences are too long.
- The main story line is in the third paragraph.
- Unexplained references: demand for what is expected to be up?

- Round numbers: not 1.637 percent.
- Elevator economics: this is up, this is down.
- Bullets preferable in the last paragraph.
- No URL link cited at the end.
- No contact or phone number provided.
- Proof read! In the last paragraph, “thought” should be “though”; “it’s” should be its” and “will be will increase” should read “to increase”.

### ***A Revised Version***

Released: September 16, 2004

#### **Consumers will spend more to heat their homes this winter**

Homeowners will pay much more this winter to heat their homes, according to the latest Heating Usage report released today by the Energy Minister. It predicts an 8% increase in spending over last winter.

Increases in prices for heating fuel are likely to generate higher spending, even in regions where demand for fuel is expected to fall. Average residential prices for natural gas are expected to be 10% higher than last winter, while household spending is expected to rise by 15%.

Tight global markets and elevated crude oil prices are expected to result in higher prices for petroleum products. The cost of imported crude oil to refineries this winter is projected to average 98 cents per gallon (about \$40 dollars per barrel), compared with 70 cents per gallon last year.

Despite above-average stocks of natural gas, average winter natural gas prices, both at the wellhead and retail levels, are expected to be above those of last winter.

Other interesting findings from this report:

- The spot price for crude oil continues to fluctuate. Prices continue to remain high even though the Organization of Petroleum Exporting Countries (OPEC) production of crude oil reached its highest levels in September since OPEC was established in 1982.
- Overall petroleum inventories are expected to be in the normal range.

See the entire report at [www.HeatingUsage.gov](http://www.HeatingUsage.gov). Contact John Smith in the Press Office at 123.4567 for more information.

## Evaluating the impact

### *Media analysis*

It is a good idea for statistical agencies to monitor the impact of their statistical stories in the print and electronic media from the point of view of both the number of “hits” and the quality of coverage. Useful resources for gauging the breadth, balance and effectiveness of media coverage include Google News, LexisNexis, blogs, and electronic and paper subscriptions.

Monitoring coverage can help managers determine if more work is needed to educate journalists, statisticians or key stakeholders about better ways of conveying the meaning of numbers in language that laypeople can understand. Monitoring would include:

- keyword searches to measure extent of media coverage;
- total coverage for a pre-determined period of time;
- daily coverage to identify spikes;
- comparing coverage to established baselines;
- prior releases of the same data product;
- qualitative methods to analyse media coverage;
- correct interpretation of the numbers;
- coverage of target audiences;
- inclusion of key story-line messages;
- inclusion of core corporate messages;
- effective use of illustrative embedded graphics;
- tone of story (positive/negative);
- tone of quotes from external spokespersons (positive/negative).

### *Website analysis*

Monitoring Internet traffic with website usage software can help determine types of stories most in demand. You should look for:

- the number of page views, visits, etc., to specific pages;
- where visitors are coming from;
- where visitors are going when they leave your pages.

In addition, surveys of users of your site – both media and general users – can help target and improve the information available. You should:

- ask the customer if they found what they were looking for when they came to the site;
- target specific questions to known users of the site;
- ask how the site is used and how often;
- assess general satisfaction with the site;
- solicit recommendations for change or additional topics;
- use focus groups with media representatives to explore needs, approaches and reactions.

## Before and after: Applying good writing techniques

To illustrate how to turn a routine statistical story into one with a much stronger story-line and more effective use of data, here is a ‘before’ and ‘after’ example. Note the differences.

### **BEFORE**

#### **Divorces – 2003**

In 2003, 70,828 couples divorced, up a slight 1.0% from the recent low of 70,155 in 2002.

The number of divorces has remained relatively stable over the last few years. The year-to-year change has been below two percent for every year since 1999.

The increase in the number of divorces between 2002 and 2003 kept pace with the increase in the Canadian population over this period. As a result, the crude divorce rate for 2003 remained the same as in 2002, at 223.7 divorces for every 100,000 people in the population.

The 1.0% increase in the number of divorces across Canada is primarily due to

a 5.1% increase in the number of divorces in Ontario and a 1.4% increase in Quebec between 2002 and 2003. Prince Edward Island and Saskatchewan were the only other provinces to experience an increase in the number of divorces between these years. Newfoundland and Labrador showed the largest percentage decrease by far in the number of divorces, down 21.4%.

Repeat divorces, involving people who had been divorced at least once before, are accounting for an increasing proportion of divorces.

In 1973, only 5.4% of divorces involved husbands who had previously been divorced. Thirty years later this proportion has tripled to 16.2% of all divorces.



The proportion of divorces involving wives who had previously been divorced is similar, rising from 5.4% to 15.7% over this thirty year period.

Marriage stability can be assessed using divorce rates based on years of marriage. The proportion of marriages expected to end in divorce by the 30th wedding anniversary inched up to 38.3% in 2003, from 37.6% in 2002.

The divorce rate varies greatly depending on how long couples have been married, rising rapidly in the first few years of marriage. The peak divorce rate in 2003 occurred after three years of marriage, when 26.2 out of 1,000 marriages ended in divorce. The risk of divorce decreased slowly for each additional year of marriage.

The custody of dependents, the vast majority of whom are children aged 18 and under, was granted through divorce court proceedings in 27% of 2003 divorces

In the remaining divorces, couples arrived at custody arrangements outside the divorce proceedings, or they did not have dependents. The number of dependents in these divorces is not available.

There has been a 17-year trend of steady increases in joint custody arrangements. Of the 33,000 dependents for which custody was determined through divorce proceedings in 2003, 43.8% were awarded to the husband and wife jointly, up 2.0% from 2002. Under a joint custody arrangement, dependents do not necessarily spend equal amounts of their time with each parent.

The custody of 47.7% of dependents was awarded to the wife and 8.3% to the husband in 2003. In 2002, these percentages were 49.5% and 8.5%, respectively.

The shelf tables Divorces, 2003 (84F0213XPB, \$22) are now available.

For general information or to order custom tabulations, contact Client Custom Services (613-951-1746; [hd-ds@statcan.ca](mailto:hd-ds@statcan.ca)). To enquire about the concepts, methods or data quality of this release, contact Brent Day (613-951-4280; [brent.day@statcan.ca](mailto:brent.day@statcan.ca)) or Patricia Tully (613-951-1759; [patricia.tully@statcan.ca](mailto:patricia.tully@statcan.ca)), Health Statistics Division.

## AFTER

### Divorces – 2003

Repeat divorces, those involving people who had been divorced at least once before, are accounting for an increasing proportion of divorces in Canada, according to new data.

In 1973, only 5.4% of divorces involved husbands who had previously been divorced. Some 30 years later, this proportion has tripled to 16.2% of all divorces. Similarly, the proportion of divorces involving wives who had previously been divorced rose from 5.4% to 15.7% during this three-decade period.

The number of couples getting a divorce in 2003 edged up 1.0% from a year earlier to 70,828. This slight increase was due primarily to a 5.1% jump in divorces in Ontario, and a 1.4% increase in Quebec. Prince Edward Island and Saskatchewan were the only other provinces to experience an advance.

The number of divorces fell 21.4% in Newfoundland and Labrador, by far the largest decline. No information on the reason for this decrease is available.

The number of divorces has remained relatively stable over the last few years.

The year-to-year change has been below 2% since 1999. The slight rise in divorces in 2003 kept pace with the increase in the Canadian population.

### Divorces

	2002	2003	2002 to 2003
	number		% change
<b>Canada</b>	<b>70,155</b>	<b>70,828</b>	<b>1.0</b>
Newfoundland and Labrador	842	662	-21.4
Prince Edward Island	258	281	8.9
Nova Scotia	1,990	1,907	-4.2
New Brunswick	1,461	1,450	-0.8
Quebec	16,499	16,738	1.4
Ontario	26,170	27,513	5.1
Manitoba	2,396	2,352	-1.8
Saskatchewan	1,959	1,992	1.7
Alberta	8,291	7,960	-4.0
British Columbia	10,125	9,820	-3.0
Yukon	90	87	-3.3
Northwest Territories	68	62	-8.8
Nunavut	6	4	-33.3

### Total divorce rate, by the 30th wedding anniversary

	2002	2003	2002 to 2003
	per 100 marriages		increase/decrease
<b>Canada</b>	<b>37.6</b>	<b>38.3</b>	<b>0.7</b>
Newfoundland and Labrador	21.8	17.1	-4.7
Prince Edward Island	25.2	27.3	2.1
Nova Scotia	30.4	28.9	-1.5
New Brunswick	27.2	27.6	0.4
Quebec	47.6	49.7	2.1
Ontario	34.9	37.0	2.1
Manitoba	30.3	30.2	-0.1
Saskatchewan	28.7	29.0	0.3
Alberta	41.9	40.0	-1.9
British Columbia	41.0	39.8	-1.2
Yukon	43.4	40.0	-3.4
Northwest Territories and Nunavut <sup>1</sup>	31.2	27.6	-3.6

1. Northwest Territories and Nunavut are combined to calculate the rates in this table because marriage and divorce data are not available for these territories separately for the 30-year period required for the calculation of the total divorce rate.

As a result, the crude divorce rate for 2003 remained stable at 223.7 divorces for every 100,000 people in the population.

Marriage stability can be assessed using divorce rates based on years of marriage. The proportion of marriages expected to end in divorce by the 30th wedding anniversary inched up to 38.3% in 2003, from 37.6% in 2002.

The divorce rate varies greatly depending on how long couples have been married. It rises rapidly in the first few years of marriage. The peak divorce rate in 2003 occurred after three years of marriage, when 26.2 out of 1,000 marriages ended in divorce.

The risk of divorce decreased slowly for each additional year of marriage.

The custody of dependents, the vast majority of whom are children aged 18 and under, was granted through divorce court proceedings in 27% of 2003 divorces.

Available on CANSIM: table 053-0002. Definitions, data sources and methods: survey number 3235.

The shelf tables Divorces, 2003 (84F0213XPB, \$22) are now available. For general information or to order custom tabulations, contact Client Custom Services (613-951-1746; hd-ds@statcan.ca). To enquire about the concepts, methods or data quality of this release, contact Brent Day (613-951-4280; brent.day@statcan.ca) or Patricia Tully (613-951-1759; patricia.tully@statcan.ca), Health Statistics Division.

## Examples of well-written statistical stories

There are many sources of well-written stories and this guide can only touch on some. You can find more examples on the Internet, in newspapers and in statistical publications. Here are a few areas to start looking:

- Statistics Norway publishes their Statistical Magazine online. It features a wide range of topics and shows examples of clear tables and graphics.<http://www.ssb.no/english/magazine/>
- The United States Bureau of Justice Statistics website links to their online publications and press releases.<http://www.ojp.usdoj.gov/bjs/>
- The United Kingdom's Office of National Statistics has a 'Virtual Bookshelf' that provides quick access to their online press releases, papers and publications, sorted by theme.<http://www.statistics.gov.uk/onlineproducts/>
- Statistics Netherlands regularly publishes short articles on the Internet as part of their 'Webmagazine' series. The articles show how to incorporate graphics to make the message clear.  
<http://www.cbs.nl/en-GB/menu/publicaties/webpublicaties/webmagazine/>
- Statistics Canada has a section on their website called 'The Daily'. Here you will find many examples of brief articles and press releases.<http://www.statcan.ca/english/dai-quo/>
- Look at websites of other statistical agencies. A good starting point is the UNECE's list of links to national and international agencies.<http://www.unece.org/stats/links.htm>

## References

Few, Stephen. (2004), *Show Me the Numbers: Designing Tables and Graphs to Enlighten*, Oakland, CA: Analytics Press.s

Kosslyn, Stephen M.(1994), *Elements of Graph Design*, New York: W. H. Freeman and Company.

Miller, Jane E., (2004), *The Chicago Guide to Writing About Numbers*, The University of Chicago Press,

Tufte, Edward R., (1997), *Visual Explanations Cheshire*, CN: Graphics Press.

Tufte, Edward R., (1990), *Envisioning information*, Cheshire, CN: Graphics Press.

Tufte, Edward R., (1983), *The Visual Display of Quantitative Information*, Cheshire, CN: Graphics Press.

Truss, Lynne, *Eats, Shoots, and Leaves: The Zero Tolerance Approach to Punctuation*, (London: Profile Books Limited, 2003)

UNECE, *Communicating with the Media: A guide for statistical organizations*, (United Nations, Geneva, 2004) <http://www.unece.org/stats/documents/media/guide/>

UNECE (2005), *Making Data Meaningful: A guide to writing stories about numbers*, Geneva

Wallgren, Anders; Wallgren, Britt; Persson, Rolf; Jorner, Ulf; and Haaland, Jan-Aage, *Graphing Statistics & Data: Creating Better Charts*, (Thousand Oaks: SAGE Publications, 1996).





# Annexes

Authors Vitæ .....	294
Abbreviations .....	303
What is DevInfo? .....	311

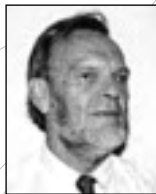
## AUTHORS VITÆ



**ADRIEN, Marie-Hélène** is President of Universal Management Group, a Canadian consulting firm specializing in evaluation and project management. She was the President of the International Development Evaluation Association (IDEAS) from 2005 to 2008. Dr. Adrién has 20 years of consulting experience in evaluation, organizational capacity development and training, representing work in 36 countries around the world. She has published a number of articles and books on evaluation, including *Organizational assessment: 25 years of lessons learned* (2005), a *Framework for improving performance* (2002), and *Enhancing organizational performance. A toolbox for self-assessment* (1999 and 2000).



**BAER, Petteri** has been working since March 2006 as Regional Adviser at the Statistical Division of the United Nations Economic Commission for Europe (UNECE) in Geneva, Switzerland. Before that he worked for 13 years in different positions at the National Statistical Office of Finland. In October 1992 he started as a marketing planner, and was soon promoted to act as Head of the Regional Services of that institution. In 2001 Mr. Baer was appointed to take the post of Marketing Manager at Statistics Finland. Mr. Baer's working career also includes elements other than statistics. He has worked in the publishing sector, as marketing manager in two publishing houses and as director of the culturally oriented bookshop "Gogol's Nose" in the city of Helsinki. In all parts of his working career he has creatively implemented new ideas and approaches to marketing.



**BAMBERGER, Michael** has a Ph.D. in Sociology from the London School of Economics. He has worked on the evaluation of development programmes in more than 30 developing countries in Africa, Asia, Latin America and the Middle East. He worked for 13 years with non-governmental organizations throughout Latin America. During his 22 years with the World Bank, he worked as advisor on monitoring and evaluation with the Urban Development Department, as Asia training coordinator for the Economic Development Institute, and as Senior sociologist in the Gender and Development Department. Since retiring from the World Bank in 2001, he has carried out consulting



and teaching assignments for the Asian Development Bank; Swedish International Development Cooperation Agency; UK Department for International Development (DFID); United Nations Development Programme (UNDP); U.N. Department of Economic and Social Affairs; UNICEF; UN Secretariat for Asia and the Pacific (ESCAP); U.N. Evaluation Office; U.S. Agency for International Development (USAID); World Bank; World Food Programme; and for several private consulting firms. Professor Bamberger has published widely on development evaluation, including a co-authored 2006 Sage publication on conducting evaluations under real-world constraints; and, several recent World Bank publications on: Conducting quality impact evaluations under budget, time and data constraints; Influential evaluations; Institutionalizing impact evaluations and reconstructing baseline data.



**FEINSTEIN, Osvaldo** is advisor to the Spanish Evaluation Agency and Professor at the Masters Degree Programme on Evaluation at the Universidad Complutense de Madrid. Professor Feinstein is a member of the monitoring and evaluation panel of the Scientific Council of CGIAR (the Consultative Group on International Agricultural Research) and evaluation consultant with the World Bank; the International Fund for Agricultural Development (IFAD); the Global Environment Fund (GEF); UNDP; CEPAL; ILPES and ILO. He was a former manager at the World Bank's Evaluation Department and former IFAD's senior evaluator responsible for Latin America, where he created PREVAL (Latin American programme for the development of evaluation capacities). Feinstein is also a former professor at the FLACSO Ecuador Master in Development Studies. He has worked in monitoring and evaluation and development in almost all Latin American and Caribbean countries, and some Asian and African countries. He has written and edited articles and books on evaluation, development and economics.



**GIOVANNINI, Enrico** graduated in Economics at "La Sapienza" University of Rome. He continued his studies at the Institute of Economic Policy of the same Faculty, specialising in econometric analysis. In December 1982 he was employed by the Italian National Institute of Statistics (Istat). In December 1989 he became research director at the National Institute for Short-term Economic Analysis, where he took care, particularly, of monetary and financial analyses. In January 1992

he moved back to the Italian National Institute of Statistics. From December 1993 to May 1997 he was head of the “National Accounting and Economic Analysis” Department. In December 1996 he was appointed Central Director of the Statistics on Institutions and Enterprises. Since January 2001 Professor Giovannini has been the Director of Statistics and Chief Statistician of OECD. He is a full professor of economic statistics at the Rome University “Tor Vergata”.



**JOBIN, Denis** is a Canadian expert in the fields of programme evaluation, performance measurement and performance audit. He is the Vice President of the International Development Evaluation Association (IDEAS) and he currently manages the evaluation unit of the National Crime Prevention Center – Department of Public Safety Canada, delivering impact evaluation studies. Mr. Jobin has been involved in evaluation-related activities for more than 13 years, having worked for the Quebec Provincial Government (Department of Industry and Trade), and the Canadian government (in the Departments of Health and Environment Canada). In addition, he has a solid experience in performance auditing, having worked for the Office of the Auditor General of Canada (OAGC). He also worked and resided in West Africa. Mr. Jobin also sponsors the Theory-based evaluation discussion group (<http://groups.yahoo.com/group/TheoryBasedEvaluation/>), and has contributed to and authored several publications related to evaluation.



**KENNEDY, Megan Grace** is a consultant with the OECD DAC Network on Development Evaluation in Paris, France, focusing on evaluation capacity development and formulating guidance on evaluating peace-building activities. Ms. Kennedy is completing a Masters of Public Administration in International Management at the Monterey Institute of International Studies in Monterey, California, USA. She holds a Bachelors degree in Economics and in Peace & Global studies from Earlham College and received a Thomas J. Watson Fellowship for independent development research in 2004. She has held various programme management positions, notably in the US, Mexico, and Tanzania.



**KHAYRI BA TALL, Oumoul** is currently the president of the International Organization for Cooperation in Evaluation (IOCE) (2008-10), past president (2005-07) and board member of the African Evaluation Association (AfrEA) and a founder of the Association Mauritanienne de Suivi-Evaluation (AMSE).

She is currently involved in initiatives to organize a network dedicated to strengthen evaluation in French speaking countries around the world (Réseau Francophone d'Evaluation, RFE). She has written several papers and articles, and delivered speeches on topics such as aid and development, and evaluation capacity. She has 21 years of professional experience in various but related field from auditing, accounting, evaluation, organisational development, micro-entreprise, micro-finance, community-based and development fields, including seven years of evaluation experience and 18 years of auditing. She is the Executive Director of her own audit and management consultancy business in Nouakchott, Mauritania. Khayri Ba Tall is an MBA (1995) and member of professional accounting bodies in Mauritania and in Senegal.



**KUSEK, Jody Zall** has provided leadership in the area of monitoring and evaluation at the World Bank for eight years. She currently heads up the Bank's Global HIV/AIDS Monitoring and Evaluation Group (GAMET) which aims to strengthen the use of HIV/AIDS data to support national and sub-national policy and programme decision-making in over 50 countries, world-wide. Previously, she was the Cluster Leader for Getting Results at the World Bank's Africa Region, and co-authored the Bank's business process to design and use a results-based country assistance strategy which is now in use, Bank-wide. Earlier, Ms. Kusek worked for the Clinton-Gore Administration in the United States, designing and implementing the Government Performance and Results Act. She is co-author of *Ten steps to results-based monitoring and evaluation*. She is also the author of numerous papers on government management, results-based management and poverty monitoring system development.



**LUNDGREN, Hans** manages the OECD/ DAC Network on Development Evaluation which brings together evaluation managers and experts from 30 bilateral and multilateral development agencies. He joined the OECD in 1987 and has since worked on development policy and aid effectiveness issues, with an increasing focus over time on development evaluation. He has published on evaluation systems in aid agencies and written a number of articles on DAC's evaluation work. He led the drafting of the DAC Principles for evaluation of development assistance and co-ordinated the work on the DAC Glossary of terms in evaluation and results-based management, and the DAC Evaluation quality standards. He also has experience from international expert reviews of monitoring and evaluation systems and is a member of UNESCO's Oversight Advisory Committee. Prior to joining the OECD in 1987, he worked in field operations with the UNDP in West Africa, and at UNESCO headquarters managing trust fund operations.



**MACKAY, Keith** is a senior evaluation officer in the Independent Evaluation Group of the World Bank, where he is also the coordinator for evaluation capacity development. His current work is focused on helping countries strengthen their national monitoring and evaluation systems to support a performance orientation within their public sectors. Countries with which he is currently working include Brazil, Chile and Colombia. Before joining the Bank in 1997, Mr. Mackay worked for 22 years in the Australian government, including 11 years in the Department of Finance. From 1991 to 1997 he was the senior adviser to the government on its national evaluation strategy. He has written 75 articles, papers and books, principally on monitoring and evaluation.



**O'BRIEN, Finbar** is Director of Evaluation at UNICEF. He has worked in international development for 25 years, fifteen of which were spent in Africa. He was formerly the Head of Evaluation and Audit with the Department of Foreign Affairs in Ireland and also served as Chair of the DAC Evaluation Network. O'Brien's major interests in recent years have been institutional arrangements for evaluation and the promotion of joint and country-led evaluations.



**OSWALT, Kris** is an international expert in the design and implementation of information systems. He has over 30 years of experience in software application development for database management systems, geographic information systems and knowledge management systems. Mr Oswald is the President of Community Systems Foundation, a not-for-profit organization founded in the USA in 1963 and the Executive Director of the DevInfo Support Group where he has been instrumental in the design of DevInfo database technology. Mr Oswald has provided technical assistance in more than 80 countries to a broad range of international organizations, including: UNICEF; UNFPA; UNDP; WFP; UN-Habitat; UNESCO; WHO; DFID; USAID; World Bank; UN Statistics Division; OECD; John Snow Inc.; International Science and Technology Institute; and, the Management Sciences for Health and U.S. Library of Congress.



**PICCIOTTO, Robert** is a graduate of the Woodrow Wilson School of Public and International Affairs (Princeton University). He is Visiting Professor at Kings College, London. He sits on the council of the United Kingdom Evaluation Society and on the board of the European Evaluation Society. At the World Bank, he served as Vice President for Corporate Planning and Budgeting and, for ten years, as Director-General, Evaluation, reporting directly to the executive directors. Prior to this, he held senior operational management assignments in three of the World Bank's regions. Since 2002, Professor Picciotto has been a senior evaluation adviser to governments and international institutions. He currently serves as a member of the International Advisory Committee on Development Impact set up by the UK Secretary of State for International Development and acts as a trustee of the Oxford Policy Institute.



**PRON, Nicolas Charles** has been working for the United Nations for 16 years, out of which 12 years were spent in the field in Africa and Asia, where he implemented UNICEF Country Programmes. Mr. Pron is currently posted in New York where he manages the DevInfo flagship project, a high profile UN inter-agency initiative to monitor progress towards achieving the Millennium Development Goals. Mr Pron is a national of France; he holds a Masters degree in International law from the Sorbonne University and a Masters degree in Development law from the Rene Descartes University in Paris.



**QUESNEL, Jean Serge** is Professor at the United Nations System Staff College, Adjunct Professor at Carleton University and Professeur Associé at the École Nationale d'Administration Publique of Quebec. He was Director of Evaluation at the United Nations Children Fund (UNICEF), the Inter-American Development Bank (IADB) and the Canadian International Development Agency (CIDA) where he was also Director of Policy Coordination, Management Improvement and International Development Programmes in Asia, Africa, Latin America and the Caribbean.



**RIST, Ray** has had a distinguished career which includes a range of high profile government and academic appointments. He has been a visiting professor at several prestigious universities, and has been a consultant to many national and international organisations, including the World Bank, OECD, DFID, IADB, and a range of corporations, and House and Senate committees in the United States. The focus of much of this consulting has been on public sector performance, especially that of results-based management and measurement, and he has been an advisor to senior government officials in more than 50 countries. Professor Rist is currently an advisor to the World Bank, co-director of the International programme for Development Evaluation Training (IPDET) and President of IDEAS. He has authored or edited 26 books and has authored more than 140 articles and monographs.



**RUGH, Jim** has been professionally involved for 44 years in rural community development in Africa, Asia, Appalachia and other parts of the world, specializing in international programme evaluation for 28 years. In 2007 he retired after serving for 12 years as head of Design, Monitoring and Evaluation for Accountability and Learning for CARE International. Rugh is recognized as a leader in evaluation among colleagues in the international NGO community, including InterAction's Evaluation Interest Group, as well as active involvement for many years in the International and Cross-Cultural Evaluation Topical Interest Group of the American Evaluation Association (AEA). He currently serves as the AEA

representative to the International Organization for Cooperation in Evaluation (IOCE). He co-authored the popular and practical *Real-World Evaluation* book and has led numerous workshops on that topic for many organizations and networks in many countries.



**SAKVARELIDZE, George** is a monitoring and evaluation specialist at the UNICEF Regional Office for CEE/CIS. He studied Pediatrics in Georgia, Tbilisi and earned a Master degree in Public Health in USA, New York at the School of Public Health in Albany. He worked with UNICEF in health and monitoring and evaluation fields. Since 2005 he has been the Regional Coordinator for Multiple Indicator Cluster Survey in CEE/CIS, coordinating 13 surveys. He also delivers technical assistance for DevInfo implementation in the Region.



**SEGONE, Marco** has been serving as the Senior regional advisor, Monitoring and Evaluation in the UNICEF Regional Office for Central and Eastern Europe and the Commonwealth of Independent States (CEE/CIS) since 2005. He represents UNICEF on the Board of Trustees of the International Programme Evaluation Network (IPEN). During his 17 years in international development, Segone has worked in Bangladesh, Pakistan, Thailand, Uganda and Albania in integrated development projects. In 1996 he joined UNICEF to work for the Regional UNICEF Office for Latin America and the Caribbean. From 1999 to 2001 he worked as Monitoring and Evaluation Officer for UNICEF Niger, where he founded, and for two years coordinated, the Niger Monitoring and Evaluation Network (ReNSE). From 2001 to 2004 he was the UNICEF Monitoring and Evaluation Officer for Brazil, where he was one of the founders and coordinator of the Brazilian Evaluation Network. In 2003 he was elected Vice-President of IOCE and was one of the founders of the Latin America and the Caribbean Network for Monitoring, Evaluation and Systematization (RELAC). Mr Segone has authored / edited about 30 books and articles, including *Bridging the gap. The role of M&E in evidence-based policy making; New trends in development evaluation, Creating and developing evaluation organizations. Lesson learned from Africa, Americas, Australasia and Europe;* and *Democratic evaluation.*



**VADNAIS, Daniel** joined UNICEF Headquarters at the end of 2006 as Data Dissemination Specialist. Prior to that, Mr. Vadnais worked for 12 years with the Demographic and Health Surveys (DHS) project as Deputy Advisor for Communication, with a focus on the dissemination of findings. He also worked closely with media representatives. Mr. Vadnais provided technical assistance in numerous countries throughout Asia and Africa. In 2006, he contributed to the publication of *Women's lives and experiences: Changes in the past 10 years*. Before that, he co-wrote *Connecting people to useful information: guidelines for effective data presentations* with members of the Dissemination working group of the MEASURE Programme. Mr. Vadnais also worked as Information officer for the Global Committee of Parliamentarians on Population and Development. In 1989 -1990, after coordinating the local arrangements of the Moscow Global Forum on Environment and Development, he served as Public Affairs Officer for Religious and Parliamentary Affairs at UNICEF/New York, at the time of the World Summit for Children. With UNICEF, he helped organize the first global inter-faith conference to focus solely on children's issues which took place at Princeton University. Mr. Vadnais, a native from Québec, holds a Masters Degree in Demography from the University of Montreal.



## ABBREVIATIONS

ADB	Asian Development Bank
AEA	American Evaluation Association
AfrEA	African Evaluation Association
CEE	Central and Eastern Europe
CIS	Commonwealth of Independent States
CES	Canadian Evaluation Society
CGD	Center for Global Development
CLE	Country-Led Evaluation
CLEF	Country-Led Evaluation Fund
CLES	Country-Led Evaluations Systems
CLIE	Country-Led Impact Evaluations
CoP	Communities of Practice
CPI	Consumer Price Index
CsOs	Civil Society Organisations
DAC-OECD	Development Assistance Committee of the Organization for Economic Cooperation and Development
DEReC	An on-line evaluation resource centre
DHS	Demographic and Health Surveys
ECD	Evaluation Capacity Development
ECG	Evaluation Cooperation Group
EGPRSP	Economic Growth and Poverty Reduction Strategy Paper
EO/UNDP	Evaluation Office of the United Nations Development Programme
EU	European Union
GBS	General Budget Support
GDP	Gross Domestic Product
IDEAS	International Development Evaluation Association
IHSN	International Household Survey Network
IOB	Dutch Ministry of Foreign Affairs
IOCE	International Organization for Cooperation in Evaluation
IPDET	International Programme for Development Evaluation Training
LFA	Logical Framework Analysis

LPA	Local Plans of Action for Children
MBO	Management by Objectives
MDGs	Millennium Development Goals
MICS	Multiple Indicator Cluster Surveys
MICS3	Multiple Indicator Cluster Surveys – third round
MfDR	Management for Development Results
MoET	Ministry of Economy and Trade
M&E	Monitoring and evaluation
MES	Malaysian Evaluation Society
NGO	Non-Government Organization
NONIE	Network of Networks for Impact Evaluation
NPA	National Plan of Action for Children
NSOs	National Statistical Offices
OECD	Organization for Economic Cooperation and Development
OECD-DAC	Development Assistance Committee of the Organization for Economic Cooperation and Development
ORET/ MILIEV	Development and Environment Related export Transactions
PRSP	Poverty Reduction Strategy Papers
QED	Quasi-experimental design
RéNSE	Réseau Nigérien de Suivi et Evaluation (Niger monitoring and evaluation network)
RWE	Real World Evaluation
SEDESOL	Mexican Secretariat for Social Development
SFE	Société Française d’Evaluation (French Evaluation Society)
SORS	Statistical Office of the Republic of Serbia
TOR	Terms of Reference
TRIPS	Trade Related Intellectual Property Rights
UNDAF	United Nations Development Assistance Framework
UNDGO	United Nations Development Group Office – now UNDOCO
UNDOCO	United Nations Development Operations Coordination Office – formerly UNDGO
UNDP	United Nations Development Programme
UNECE	United Nations Economic Commission for Europe

UNEG	United Nations Evaluation Group
UNFPA	United Nations Population Fund
UNGASS	United Nations General Assembly Special Session
UNICEF	United Nations Children's Fund
UNICEF IRC	UNICEF Innocenti Research Center
3ie	International Initiative for Impact Evaluation













## WHAT IS DEVINFO?

DevInfo is a powerful database system which monitors progress towards the Millennium Development Goals and Human Development. It generates tables, graphs and maps for reports and presentations. DevInfo has been developed by United Nations organizations. It was adapted from UNICEF ChildInfo technology. The database maintains indicators, by time periods and geographical areas, to monitor commitments to sustained human development.

UNICEF Regional Office for Eastern and Central Europe and the Commonwealth of Independent States developed three regional databases. The Regional MDGInfo database, developed in cooperation with UNECE and UNDP, makes MDGs as well as regionally specific indicators easily available. It is accessible at [www.regionalmdg.org](http://www.regionalmdg.org). The MICSInfo database presents the key findings of the third round of Multiple Indicators Clusters Surveys carried out in 12 countries in the region, with data disaggregated by regions, urban and rural, ethnicities, wealth quintiles, mother's education and age of children. It is accessible at [www.micsinfo.org](http://www.micsinfo.org). Last but not least, the MoneeInfo database makes data on the situation of children and women, with a specific focus on child protection, easily accessible at [www.moneeinfo.org](http://www.moneeinfo.org).

All three databases are now available in the CD ROM attached to this report. In the CD ROM, you can also download ready-made graphs and maps on key indicators, the full database in Excel format and produce your own maps, graphs and table using the DevInfo technology.

For additional information on DevInfo, and a quick guide on how to produce maps, graphs and tables using the DevInfo technology, please visit [www.devinfo.org](http://www.devinfo.org).

### **Instructions on installation and use of DevInfo**

Ready-made graphs and maps on the key indicators, as well as the full database in Excel format, are accessible immediately. To produce your own maps, graphs and table using the DevInfo technology, you need to install DevInfo in your computer. Below the instructions.

## ***System requirements for DevInfo***

The recommend hardware requirements to install this software application are:

- Pentium IV
- 512 MB of RAM
- 1 GB of free hard disk
- Display resolution 1024 x 768
- Microsoft Windows XP
- Microsoft Office XP
- Microsoft Internet Explorer 6.0

## ***Installing DevInfo***

To install this software application on your computer, follow the steps given below:

- Insert the CD ROM into the drive of your computer
- Wait for auto-run to open the setup screen
- Click on the icon “Database”
- Follow the instructions on screen to complete the setup
- Double-click on the DevInfo icon on desktop to start the application

### ***If the setup program does not load automatically:***

- Choose Start | Run
- Type d:\setup where d is the letter of your CD-ROM drive and press Enter key
- Follow the instructions on screen to complete setup
- Double-click on the DevInfo icon on desktop to start the application

Note: Computers with Windows 98 Operating System need to be restarted after installing DevInfo.

---

UNICEF Regional Office for CEE/CIS  
Palais des Nations  
CH 1211 Geneva 10  
Switzerland  
[www.unicef.org/ceecis](http://www.unicef.org/ceecis)

2009