giz

Better Monitoring for Better

Evaluation

Draft Version of Paper Document

Wednesday, 06 April 2011

Content

Better Monitoring for Better Evaluation

Module 1: Introduction: better monitoring for better evaluation

Contents

1. Course concept

1.1.   Aims of the course

The Purpose of this course is to improve the monitoring skills of key stakeholders involved with poverty reduction and other development strategies.

The Goal of the course is to improve the quality of monitoring of PRSPs and other development strategies, and to thereby improve the information base available for the evaluation of impact and outcomes.

The target audience for training is initially ECOWAS and its member states.

1.2.   Paris Declaration and Accra Agenda

The foundations of this course lie in the Paris Declaration on Aid Effectiveness (2005) and the subsequent Accra Agenda for Action (2008). From the Paris Declaration, we highlight the following extracts:

3.   We reaffirm the commitments made at Rome to harmonise and align aid delivery. We are encouraged that many donors and partner countries are making aid effectiveness a high priority, and we reaffirm our commitment to accelerate progress in implementation, especially in the following areas:

i. Strengthening partner countries' national development strategies and associated operational frameworks (*e.g.*, planning, budget, and performance assessment frameworks).

4. We commit ourselves to taking concrete and effective action to address the remaining challenges, including:

i. Weaknesses in partner countries' institutional capacities to develop and implement results-driven national development strategies.

Managing resources and improving decision-making for results

43. Managing for results means managing and implementing aid in a way that focuses on the desired results and uses information to improve decision-making.

44. Partner countries commit to:

• Endeavour to establish results-oriented reporting and assessment frameworks that monitor progress against key dimensions of the national and sector development strategies; and that these frameworks should track a manageable number of indicators for which data are cost-effectively available (Indicator 11).

45. **Donors** commit to:

• Link country programming and resources to results and align them with effective partner country performance assessment frameworks, refraining from requesting the introduction of performance indicators that are not consistent with partners' national development strategies.

• Work with partner countries to rely, as far as possible, on partner countries' results-oriented reporting and monitoring frameworks.

• Harmonise their monitoring and reporting requirements, and, until they can rely more extensively on partner countries' statistical, monitoring and evaluation systems, with partner countries to the maximum extent possible on joint formats for periodic reporting.

We note the main thrust which is to minimise the number of development monitoring systems which are operational within a country and a wish to rely on national systems wherever possible. This must of necessity imply better quality national systems and more reliance on the standards outputs of government systems.

Three years later in Accra, some progress had been made but there was still cause for concern in a number of areas, as the following extract shows:

23. We will improve our management for results by taking the following actions:

a) Developing countries will strengthen the quality of policy design, implementation and assessment by improving information systems, including, as appropriate, disaggregating data by sex, region and socioeconomic status.

b) Developing countries and donors will work to develop cost-effective results management instruments to assess the impact of development policies and adjust them as necessary. We will better co-ordinate and link the various sources of information, including national statistical systems, budgeting, planning, monitoring and country-led evaluations of policy performance.

c) Donors will align their monitoring with country information systems. They will support, and invest in strengthening, developing countries' national statistical capacity and information systems, including those for managing aid.

Our course will support these aims.

1.3. Rationale of the course

As the extracts from the Paris Declaration and Accra Agenda show, the spotlight continues to shine on more effective and more efficient monitoring to ensure that the desired development goals are being met.

But despite numerous initiatives proposing and supporting systems to monitor development programmes, plans and strategies within the ECOWAS region, it is still felt that the people most involved with implementing the monitoring systems require additional training. One of the problems is that many monitoring and evaluation systems have been set up by a limited group of experts with not enough attention paid to national needs and existing data, or the scarcity of resources and suitably trained manpower.

It is true that national strategies will demand very different approaches, in particular taking into account the size of the country and its administrative organisation. There is a high degree of regional variability, and this puts ECOWAS in a challenging position to monitor on-going and planned processes, compile comparable results and statistics, and counsel and advocate on solutions which are regionally relevant.

An added complication is the plethora of programmes and sectoral strategies, which, though normally operating under the umbrella of a comprehensive PRSP, in reality demand separate and sometimes contradictory monitoring arrangements.

Our aim therefore is to focus on those people in member states and ECOWAS who are involved in regular monitoring of plans and programmes, and to give them a better

understanding of issues around monitoring and evaluation. A key part of the course will be linking data producers, including the national statistical service, with users.

1.4. Stakeholders in the course

It is clear that a stakeholder meeting will be needed to identify the needs, the requirements and the optimal composition of the stakeholder group for the various types of trainings before they start.

On a regional level we will address the different ECOWAS commissions dealing with subject matters related to the commission, for example the Macro-Economic Policy Commission dealing with issues such as Private Sector promotion. It is expected that the Research and Statistics Department will be among the stakeholders at regional level.

On a national level various sector planners should be included as well the data producers, whether these are in the Statistical Services or line ministries. It seems important to address those stakeholders responsible for monitoring the current strategies and plans. It is likely that middle level staff members of these organisations will be the appropriate target group.

1.5. Problem analysis

A key element of our training is the problem analysis. Various requests and references to existing problems show a persistent demand for project management skills. The strategic plan or a poverty reduction strategy is a project like many others; and a log frame approach to identify the above mentioned steps might help. At first the problems of monitoring a plan or strategy might not even been recognised, as long as the plan is set up, approved and a

summary given at the end of the period. This has to do with not addressing monitoring and evaluation needs in the design of the plan or strategy.

More than that, continuous and consistent monitoring requires additional effort and organisational skills, and also determination and an awareness of accountability. It is assumed that if the plan administrator is responsible for its outcomes then he/she will put some effort into the monitoring of success or what impeded success and at which stage.

Important information to be gathered prior to the course being run will be the past activities on monitoring and evaluation – what has been done, what has been left out, where we can build on past experience, where can we expect synergies and mutual learning. Here external factors preventing project management and success can be extremely important, but they have to be identified.

A crucial part will be to identify the information sources and various information techniques (quantitative statistical or qualitative opinion based) and the frequencies with which this information will be expected, setting milestones and identifying quality assurance methods to maintain the achievement of the objectives or record the deviations from the expected goals.

1.6. Guiding principles

Underlying the course are a number of key principles.

First, the course will focus on monitoring rather than on evaluation. We believe that linking monitoring and evaluation together as "M&E" is not helpful. Monitoring and evaluation require different skills, different types of data, and are usually done by different people. Specifically, monitoring involves greater numbers of people and is more immediate.

Second, we make no distinction between *results-based monitoring* and *monitoring*. We feel that results-based monitoring has grown up as a response to too much focus (by monitors) on the activity and output levels of a project. In reality, monitoring takes place at all levels of a project including results and outcomes.

Third, we shall follow accepted international guidelines and definitions.

Fourth, we shall spend some time on logical frameworks. As stated in the UNDP Handbook on Monitoring and Evaluation (2009):

Without proper planning and clear articulation of intended results, it is not clear what should be monitored and how; hence monitoring cannot be done well.

We believe that the logical framework approach is a good way to encourage clear articulation of results. If the *words* are right, then it is clearer what should be *monitored*. And if the *monitoring* is done well, the *evaluation* will be more efficient.

Fifth, the course aims to strengthen the links between development practitioners and the national statistical system, and should also provide information to feed into the national statistical development strategy.

2. Summary of the contents of modules 2 to 9

2.1. Module 2: Monitoring: international concepts and best practice

2.2. Module 3: Logical analysis: a way to improve monitoring

2.3. Module 4: Case studies (1): designing logframes to improve monitoring

2.4. Module 5: Indicators: selecting and using them for better monitoring

2.5. Module 6: Poverty: measuring and monitoring it for development interventions

2.6. Module 7: Case studies (2): choosing indicators for effective monitoring

2.7. Module 8: Monitoring systems: setting them up and using them for improved decision-making

2.8. Module 9: Evaluation: assuring efficient evaluation based on quality monitoring

3. Suggested preparation and pre-course reading

3.1. *Pre-course preparation*

3.2. *Key documents for pre-course reading*

3.3. *Key websites for pre-course familiarisation*

Module 2: Monitoring: international concepts and best practice using information

Content

1. Introduction

1.1. Aim and Learning goals

The aim of this module is to create an awareness of some of the dominant organization in the area of development cooperation and look at a typical of their contributions to monitoring and/or evaluation of development strategies. Participants should be able to learn from these practises.

Participants should refer to best practices of international organization. They should also limits to adapt to their national or regional needs. Some insight about the sometimes hidden agendas of development cooperation should be gained. Time permitting, exercises should allow participant to design monitoring and evaluation schemes with reference to the strategies exposed in this module

1.2. General considerations

A course for M&E does not want to reinvent the wheel or convince the auditorium there will something completely new coming up in this course. This will be about the last time (at least in this chapter the acronym M&E was used, because the two aspects or phases Monitoring and Evaluation are as naturally and inseparably linked as it sometimes might seem. It will be, not surprising, old wine in new bottles but these bottles should be appropriate for the circumstances and tasks in view. Development Plans can be complex subjects; however the principles of monitoring such plans will not be something completely unique and new. And DP have been designed before, the have been executed. Sometimes they even have been monitored and usually a plan at the end of its span of life is given an appraisal and an evaluation, if it has been a success or it has failed. More often we will come across

successful plans because the designers and implementers of past plans would expect to be summoned to design and implement the next plan.

Comparing development plan with business plans can quite enlightening although, it is agreed that the two: social / economical development and business cannot be measured though the same lenses. The very essence of business usually is to make revenue; however this may only be true on the short run but may not pay off on the long run. Even business does not stand alone in making money, responsible business success is paired with sustainable economical success, defendable business ethics, environmental and social responsibilities. This may be comparable to social development plans because most of these long-term criteria will apply as well. There is however a striking difference in measuring business success and comparing it to political or social development. Business success can be measure immediately in money terms, the success of a development plan can only be measured according to an agreed set of criteria the goals, objectives or tasks.

Without going into detail now, there are many examples of strategies proposed and defended by International Agencies. We will try to separate them into useful references for being able to apply best practices for concrete examples.

1.3. *International Agencies, Institutes and Development Plans*

If it comes to this, there are always the same suspects and before going to their portfolio and their characteristics, let us recall that all international agencies have an agenda, and more than often this agenda differs from their intentions. This is not a statement of mistrust but a simple reality of economical and political life. If you doubt this then refer to [1]. Without going much into detail, this does not mean to avoid learning from the vast experience these organizations have, stored and distributed for the benefit of members, associates and customers.

Again we will not go so much into detail because the purpose of this module is that of a reference of best practice.

Let us look at agencies, institutes, organizations being involved in development strategies and will give abundant material having selected four will give abundant material for referring to the strategic approach of these four and look at the best practices and the relative reports not without the awareness that non of these reports will focus on any negative aspects. Another approach of a successful business approach will be included to complete the picture and compare it briefly with the development approaches. This will be **software project management** and similarities; discrepancies will be shown as well as possible synergies among the project (since a development plan is also a project, albeit a complex one) can be found.

2. UNDP – What is it and why United Nations Development Fund

Results-based management provides a set of principles, approaches, and tools which can help us achieve the Millenium Development goals. By always trying to answer the "so what difference does our intervention make?" question, we will keep our focus on how we can support real and sustainable improvements being made in the lives of those we serve.

This, in turn, requires us to embrace a culture of evaluation. The tremendous store of knowledge which UNDP has, and has access to, including the lessons learned from evaluations, should more fully inform our programming and our decision making.

This initial remarks form [11] should be enough to justify a look at the proposals and experiences of UNDP.

The updated Handbook (2009) marks a departure from the previous publication. First, recognizing that results planning is a prerequisite for effective programme design, monitoring

Module 2: Monitoring: international concepts and best practice using information

and evaluation, the revised Handbook integrates planning, monitoring and evaluation in a single guide. Second, the Handbook reflects the requirements and guiding principles of the evaluation policy, including national ownership, which is now mainstreamed throughout the cycle of planning, monitoring and evaluation. Third, the revised Handbook includes a comprehensive chapter on evaluation design for quality assurance. Even though it is mentioned in third place, the inclusion of QA is very important and we will have an intense look at the approach.

2.1. *Planning for Monitoring and Evaluation*

Monitoring and evaluation serve several purposes. In the absence of effective monitoring and evaluation, it would be difficult to know whether the intended results are being achieved as planned, what corrective action may be needed to ensure delivery of the intended results, and whether initiatives are making positive contributions towards human development. Monitoring and evaluation always relate to pre-identified results in the development plan. They are driven by the need to account for the achievement of intended results and provide a fact base to inform corrective decision making.

(We highlighted the important part in these sentences)

Let us briefly gave a look at the Result Based Monitoring terminology (not very different from

The results framework					
Results	**Indicators**	**Baseline**	**Target**	**Means of Verification**	**Risks & Assumptions**
Impact statement (*Ultimate benefits for target population*)	Measure of progress against impact				Assumptions made from outcome to impact. Risks that impact will not be achieved.
Outcome statement (*Short- to medium-term change in development situation*)	Measure of progress against outcome				Assumptions made from outputs to outcome. Risks that outcome will not be achieved.
Outputs (*Products and services—tangible and intangible— delivered or provided*)	Measure of progress against output				Assumptions made from activities to outputs. Risks that outputs may not be produced.
Activities (*Tasks undertaken in order to produce research outputs*)	Milestones or key targets for production of outputs				Preconditions for implementation of activities.

OECD and WB because these organizations are closely linked)

Here again we find the (four against five in other units) different layers and the planning

process is geared around this framework. The results framework can be completed with all

the outcomes, outputs, activities and inputs that the stakeholders have identified. However,

in many cases, a more limited framework showing only the specific outcomes and outputs

related to a particular agency (such as UNDP) and its partners will be needed to satisfy

internal requirements. In these cases where a more focused results framework is created,

every effort should be made to show information on the broader agenda of actions being

pursued and the partners and non-partners working towards achieving the overall outcomes

Module 2: Monitoring: international concepts and best practice using information

and impacts in the narrative of the wider strategy document. The strategy document should

not be confined to only what the agency will produce. It should instead show how the efforts

of different stakeholders will contribute to the achievement of the common overall vision and

intended impacts. This will also aid the monitoring and evaluation processes.
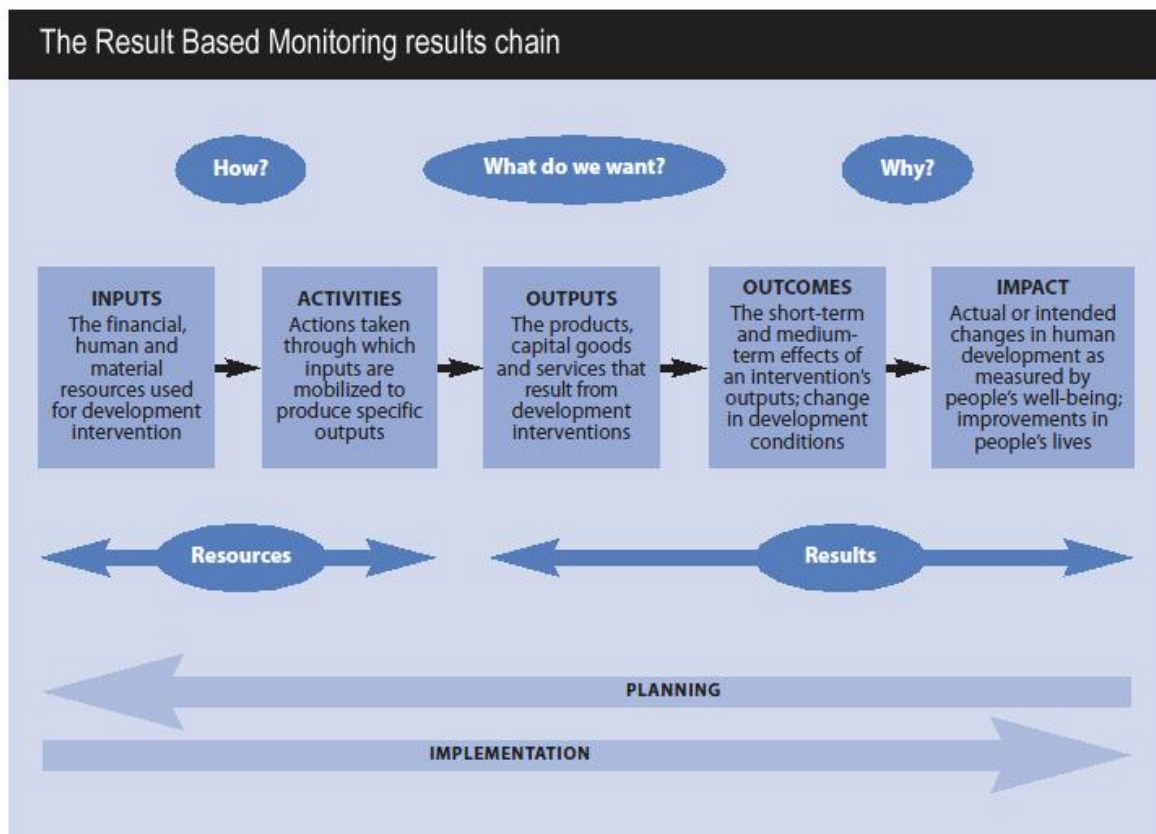
Here we find again the five layers of the planning process and nicely displayed the distinction

between implementation (here resources) and results. As a result based planning the

process starts from the most complex goal, the impacts and the implementation will start

from the most elementary: the inputs. The Monitoring and Evaluation process will be closely

linked

2.2.   *Monitoring and Evaluation Framework*

A clear framework, agreed among the key stakeholders at the end of the planning stage, is

essential in order to carry out monitoring and evaluation systematically. This framework

serves as a plan for monitoring and evaluation, and should clarify:

What is to be monitored and evaluated

The activities needed to monitor and evaluate

Who is responsible for monitoring and evaluation activities

When monitoring and evaluation activities are planned (timing)

How monitoring and evaluation are carried out (methods)

What resources are required and where they are committed

What will be needed and used as proposed by UNDP is a Planning matrix for monitoring

(and evaluation) as you see below. You can look at detailed examples of monitoring and

evaluation planning matrix at [12] and [13].

Planning matrix for monitoring							
Expected Results (Outcomes & Outputs)	Indicators (with Baselines & Indicative Targets) and Other Key Areas to Monitor	M&E Event with Data Collection Methods	Time or Schedule and Frequency	Responsibilities	Means of Verification: Data Source and Type	Resources	Risks
Obtained from development plan and results framework.	From results framework. Indicators should also capture key priorities such as capacity development and gender. In addition, other key areas need to be monitored, such as the risks identified in the planning stage as well as other key management needs.	How is data to be obtained? Example: through a survey, a review or stakeholder meeting, etc.	Level of detail that can be included would depend on the practical needs. In UNDP, this information can also be captured in the Project Monitoring Schedule Plan from Atlas.	Who is responsible for organizing the data collection and verifying data quality and source?	Systematic source and location where you would find the identified and necessary data such as a national institute or DevInfo.	Estimate of resources required and committed for carrying out planned monitoring activities.	What are the risks and assumptions for carrying out the planned monitoring activities? How may these affect the planned monitoring events and quality of data?

3. WB – What is it and why World Bank

Module 2: Monitoring: international concepts and best practice using information

How is the WB involved in development? The term "World Bank" generally refers to the International Bank for Reconstruction and Development (IBRD) and International Development Association (IDA), whereas the World Bank Group is used to refer to the institutions collectively. If we look at the Focus of the WB it reads:

A World Free of Poverty

The World Bank's projects and operations are designed to support low-income and middle-income countries' poverty reduction strategies. Countries develop strategies around a range of reforms and investments likely to improve people's lives from universal education to passable roads, from quality health care to improved governance and inclusive economic growth. In parallel, the Bank strives to align its assistance with the country's priorities and harmonize its aid program with other agencies to boost aid effectiveness.

The World Bank also strives to tackle global challenges from international trade to climate change and debt relief. It does so within each country's specific socio-economic context, adapting programs to country capacity and needs. (source:web.worldbank.org)

3.1. What is important in the context of this module?

The Independent Evaluation Group (IEG) (formerly known as the Operations Evaluation Department (OED)) plays an important check and balance role in the World Bank.

Impact evaluations: In recent years there has been an increased focus on measuring results of World Bank development assistance through impact evaluations. An impact evaluation assesses the changes in the well-being of individuals that can be attributed to a particular project, program or policy. Impact evaluations demand a substantial amount of information, time and resources. Therefore, it is important to select carefully the public actions that will be evaluated

3.2. *Controversial appraisal of the WB*

Just to cite one of many: The World Bank has long been criticized by a range of non-governmental organizations and academics, notably including its former Chief Economist Joseph Stiglitz, who is equally critical of the International Monetary Fund, the US Treasury Department, and US and other developed country trade negotiators.. Critics argue that the so-called free market reform policies – which the Bank advocates in many cases – in practice are often harmful to economic development if implemented badly, too quickly ("shock therapy"), in the wrong sequence, or in very weak, uncompetitive economies. World Bank loan agreements can also force procurements of goods and services at uncompetitive, non free-market, prices [5] (source: en.wikipedia.org)

Since its creation in 1973, the World Bank's Independent Evaluation Group (IEG)—formerly known as the Operations Evaluation Department— has supported the efforts of governments in developing countries to strengthen their monitoring and evaluation (M&E) systems and capacities. Over this period, IEG has accumulated considerable experience in this topic in a wide range of countries and public sector environments. In 1998, IEG initiated a series of working papers and other publications on this topic to document and to help disseminate these lessons.

3.3. *Structural adjustment Programs (SAP)*

Structural adjustments are the policies implemented by the International Monetary Fund (IMF) and the World Bank (the Bretton Woods Institutions) in developing countries, mainly in the 80. These policy changes were conditions for getting new loans from the IMF or World Bank, or for obtaining lower interest rates on existing loans. The Structural Adjustment Programs were created with the goal of reducing the borrowing country's fiscal imbalances. The bank from which a borrowing country receives its loan depends upon the type of necessity. The SAPs were supposed to allow the economies of the developing countries to become more market oriented. This then was supposed to force them to concentrate more on trade and production to boost their economy.

Through conditions, Structural Adjustment Programs generally implement "free market" programs and policy. These programs include internal changes (notably privatization and deregulation) as well as external ones, especially the reduction of trade barriers. Countries which fail to enact these programs may be subject to severe fiscal discipline. Critics argue that financial threats to poor countries amount to blackmail; that poor nations have no choice but to comply.

Since the late 1990s, some proponents of structural adjustment such as the World Bank, have spoken of "poverty reduction" as a goal. Structural Adjustment Programs were often criticized for implementing generic free market policy, as well as the lack of involvement from the country. These PRSPs essentially take the place of the SAPs. Some believe that the increase of the local government's participation in creating the policy will lead to greater ownership of the loan programs, thus better fiscal policy.

3.4. Example: How to Build M&E Systems to support Better Government

The example is not one; it summarizes various elements of a "good" M&E system and presents successful country reports. To our understanding it would be equally important to present unsuccessful or non sustainable attempts to learn which approaches led to failure but this would mean to criticize past methodologies like SAPs which is unlikely from the side of the authors. Apart from this there are various sections were lessons can be learned and it will be the goal of this unit to do so.

3.5. How to establish or strengthen a Monitoring and/or Evaluation system

One of the lessons to heed in building an system for Monitoring and/or Evaluation is the importance of conducting a **country diagnosis**. This provides a sound understanding of

current efforts because in most cases there is already an abundant set of monitoring activities. A diagnosis helps focus key stakeholders within government and in the donor community on the strengths and weaknesses of current arrangements in the government. A diagnosis naturally leads to an action plan that identifies the main options for strengthening a government system of Monitoring and/or Evaluation [6]. The most important lesson could be a harmonization of an existing monitoring system.

3.6. What can be learned?

Often parallel systems exist depending mostly on the state of governmental organisation: So usually in industrialized countries a system exits, a set of performance indicators, rapid reviews, impact evaluations, and performance audits. Some countries have succeeded in building a whole-of-government Monitoring and/or Evaluation system. Others, such as Uganda, use a largely uncoordinated and disparate collection of about 16 separate sector monitoring systems. If a bottom up approach is possible (Existing Regional Monitoring systems), then this can be exploited. More often these systems are not coordinated and linked to specific projects. In the second case a regional system should be used harmonizing the existing tools. In general a regional approach has to be emphasized. Without knowing what is going on a regional level (provincial, district, village) no national planning and development makes sense.

3.7. From diagnosis to action plan

A central theme is that there is no "best" model of what a government or sectoral Monitoring and/or Evaluation system should look like. Countries differ substantially in the emphasis they choose to give to the different purposes of an Monitoring and/or Evaluation system. And each country is unique in terms of its Monitoring and/or Evaluation activities, functions, and systems and in terms of its public sector culture and environment. Thus, an action plan to

create or strengthen an existing government system for Monitoring and/or Evaluation has to be tailored closely to country circumstances. This is why it is so important to conduct a diagnosis of a country's Monitoring and/or Evaluation strengths and weaknesses as a basis for developing an action plan.

3.8. Vision for the Future System

A diagnosis might ideally lead to a clear statement of what the Monitoring and/or Evaluation system would desirably look like at some time in the future. The dimensions of the system— its architecture—would include, in particular, the following elements:

1. Which of the four possible uses of Monitoring and/or Evaluation information will be the objectives of the system, recognizing the trade-offs between them: to support budget decision making or national planning; to help ministries in policy formulation, policy analysis, and program development; to support ministries and agencies in managing their activities; or to strengthen accountability relationships

2. The formal roles and responsibilities of the key stakeholders of the Monitoring and/or Evaluation system—those who would be expected to produce monitoring information and evaluations and to make use of them. It is important to locate responsibility for management or oversight in a powerful ministry, committee, or other entity.

3. Whether a whole-of-government system, including all central ministries, sector ministries, and agencies, is to be developed or if the system is to be more narrowly focused on individual sectors or agencies. Some governments develop an Monitoring and/or Evaluation system only for donor-funded operations in their country.

4. The levels of government at which the Monitoring and/or Evaluation system will be developed: central, state/provincial, or local.

Once the vision of the fully functioning Monitoring and/or Evaluation system has been formulated, it is possible to develop an action plan to achieve it. This action plan should draw on the international lessons from building country Monitoring and/or Evaluation systems [6,

chapter 10]. Of course, it would not make sense to develop an action plan unless there is already some substantive demand for Monitoring and/or Evaluation within the government.

3.9. Some Tools, Methods and Approaches for Monitoring and Evaluation

The WB Group summarizes a number of tools [8] :

We will not go much further her but look at the following graph which bears several similarities to the OECD scheme (bigger in [7]):
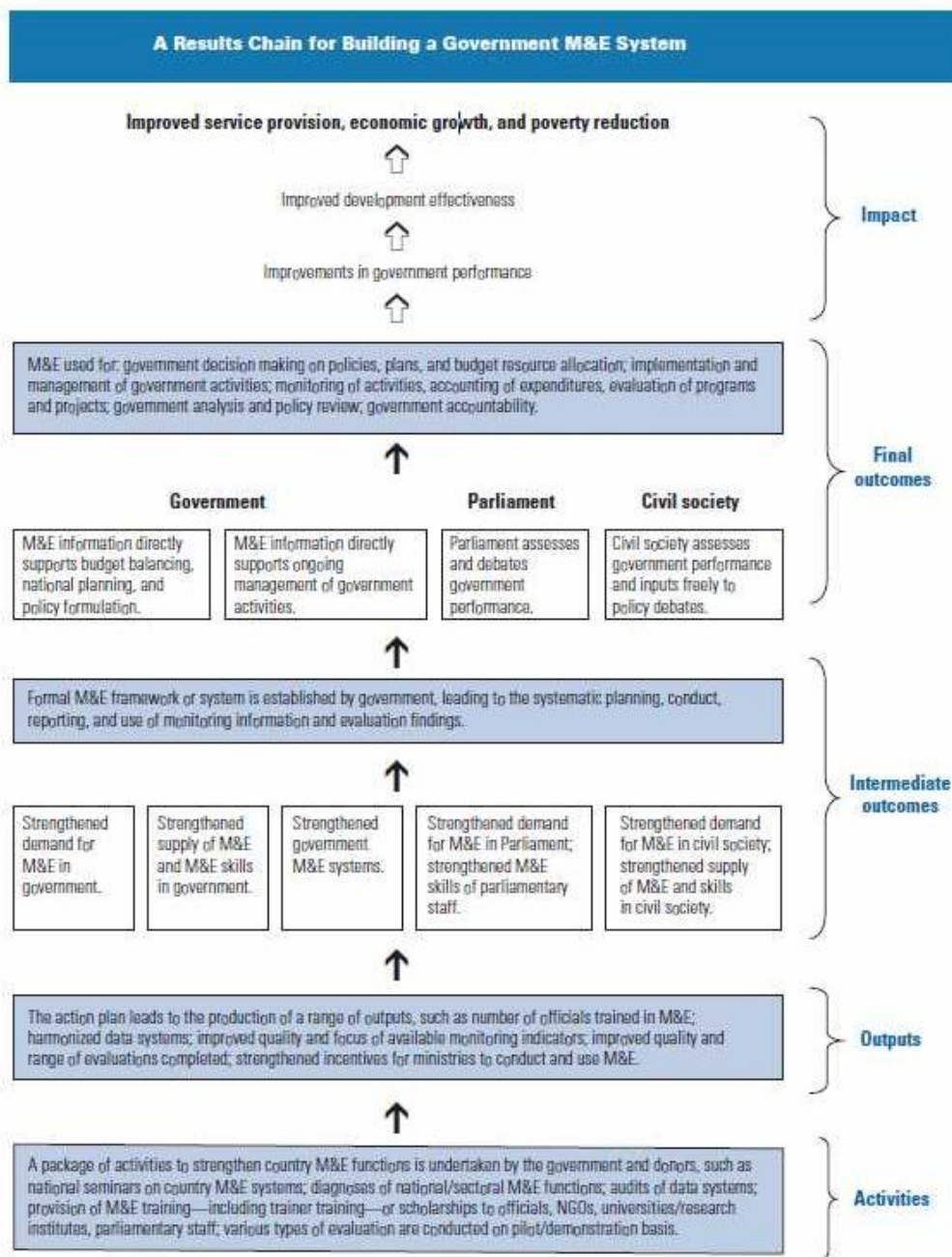
structure is prepared for these. Let us look at some techniques and tools.

The M&E Overview [8] discusses:

Performance indicators

The logical framework approach

Theory-based evaluation

Formal surveys

Rapid appraisal methods

Participatory methods

Public expenditure tracking surveys

Cost-benefit and cost-effectiveness analysis

Impact evaluation

This list is not comprehensive, nor is it intended to be. Some of these tools and approaches are complementary; some are substitutes.

Since many of these tools / techniques will be discussed and applied at another Module / Unit there will be no detailed discussion, except another almost contradicting publication of the same year of the WB Group: Building a Results-Based Monitoring and Evaluation System

3.10. Results-Based Monitoring

This document [9] is postulating that Results-Based Monitoring and / or Evaluation which may be conducted at the project program or policy level are superior to Implementation Monitoring. Traditional implementation-focused systems are designed to address compliance. Did they mobilize the needed inputs? Did they undertake and complete the agreed activities? Did they deliver the intended outputs (the products or services to be produced)? The implementation approach focuses on monitoring and assessing how well a project, program, or policy is being executed.

However, this approach does not provide policymakers, managers, and stakeholders with an understanding of the success or failure of that project, program, or policy. Results-based systems are designed to address the fact that outputs have been generated?

Results-based systems help answer the following questions

• What are the goals of the organization?

• Are they being achieved?

• How can achievement be proven?

In reality there is no contradiction: a Monitoring on various levels the OECD or the WB scheme can provide enough comprehensive information to answer all these questions but at different layers of the planning process, by different means and by different people.

3.10.1.   Translating Outcomes into Outcome Indicators

When we consider measuring "results," we mean measuring outcomes, rather than only inputs and outputs. However, we must translate these outcomes into a set of measurable performance indicators. This certainly is true but looking at the steps proposed in [9]
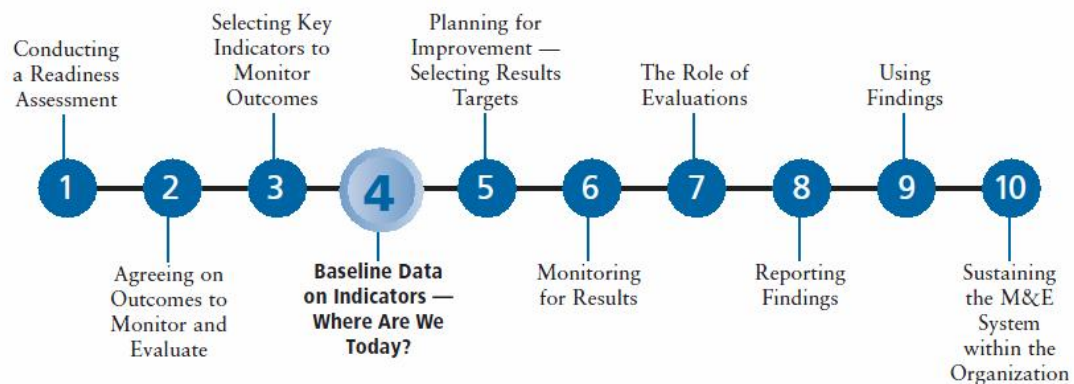
it raises the impression that there is a linear approach to a "better and result based" monitoring and evaluation. However, the discussion on baseline data around the First target of the Millennium Goals: "Halve until 2105 the number of people whose income is less than 1 US /day" raises the awareness how change / progress can be measured and on the importance of indicators. But the issues of indicators are dealt with in detail in another module of this course

3.10.2. Conclusion

Results-Based Monitoring

Results
- Goal (impacts) — • Long-term, widespread improvement in society
- Outcomes — • Intermediate effects of outputs on clients

Implementation
- Outputs — • Products and services produced
- Activities — • Tasks personnel undertake to transform inputs to outputs
- Inputs — • Financial, human, and material resources
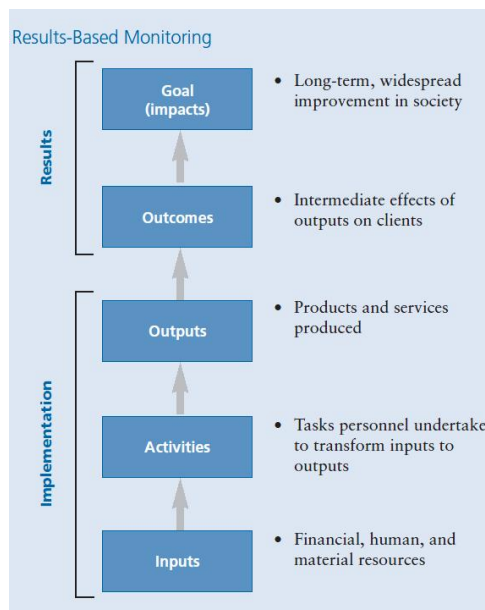
A graph in [9] shows quite clearly that there is no antagonism among implementation or result based monitoring. It happens on different stages of the planning process. Helpful in the mentioned documents is to show the various consecutive steps needed to observe the results of a planning process. Here the linear approach is helpful to undermine the idea of a check-list of necessary preconditions and essential bases for a sound system. Likewise country examples lead the way to establish a Monitoring and/or Evaluation system. All references should be combined and as central theorem is should be remembered:

There is not only way to a Monitoring and / or Evaluation system. There are frameworks to help establishing such a system and that the industrialized countries are not that much advanced establishing a monitoring system to reduce risks on vital issues shows the financial crisis and near economical collapse of the world financial system with no prior warning generated.

4. GIZ – What is it and why GIZ

The Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) is all about established on 1 January 2011, it brings together under one roof the long-standing expertise of the Deutscher Entwicklungsdienst (DED) gGmbH (German Development Service), the Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH (German Technical Cooperation) and Inwent – Capacity Building International, Germany. GIZ supports the German

Module 2: Monitoring: international concepts and best practice using information

Government in achieving its objectives in the field of international cooperation for sustainable development.

4.1. GIZ approach to Monitoring and / or Evaluation

First we will look at the classical approach being pursued for many years by the GTZ, one of the elements of the merger to become GIZ. It is called ZOPP. For many years the acronym ZOPP has stood for Objectives-oriented Project Planning. It has become GTZ's trademark for participative planning procedures geared to the needs of partners and target groups. It is something very similar to the before mentioned "Results-Based Monitoring" although different organizations tend to think of a very unique approach.

Quite recently another approach can be observed at GIZ, it is called "Systemic evaluation" and it appeared some ten years ago as a holistic approach to evaluate a whole "system" instead of analyzing a cause-effect chain. This approach will only be mentioned briefly but it seems to have sufficient ground at GIZ. It has close relation to psycho-analytical approaches by Systemic Constellation approach by German psychotherapist Bert Hellinger and Transactional Analysis by Eric Berne. It seems doubtful that is can gain acceptance on a more technical minded audience on international level.

4.2. Objectives-oriented Project Planning (ZOPP) by GIZ

In recent years (since 2000) the purely technical approach of ZOPP has been blended with Participation – a strategic orientation of programme-based approaches a modification certainly due to the lengthy discussion on ownership of any present and future international cooperation.

Let us have a look at the characteristics [10]: Projects can only be successful if target groups and partner organisations accept them and are actively committed to achieving the agreed development status. Figuratively speaking, a project planning process takes the "bottom- up" direction. It begins with the declaration by the target groups on what their needs and goals are, and the project is generated from this. Nevertheless, the chief components of a project's hierarchy of goals have often been set before project planning begins. The development-policy goals are prescribed from outside. The participating organisations operate along set guidelines even though the actual overall goal may only be detailed during the course of project planning. From the planner's viewpoint, the development-policy goals and the overall goal have been established "top-down".

Strategy	Indicators	Assumptions	Indicators of the assumptions
Overall goal: Superior strategic goal for the project			
Development goal: The changed situation designed by the target groups	How to recognise whether the development goal has been achieved		
Project purpose: Change in actions of the users of the project's services	How to recognise that the project purpose has been achieved	Matters outside the influence of the target groups which must happen for them to achieve their development goal	How to recognise that the assumption has taken place
Results: Products and services generated by the project management	Major characteristics of the results	Matters outside the project which must happen if the project purpose is to be achieved	How to recognise that the assumption has taken place
Activities to achieve the results	Quantities and costs		

Project Planning Matrix

Here we have a verbal description of the various layers of the planning process as in the previous units.

An interesting tool to support the Planning process and monitor its success is the Project Planning Matrix: Although the left two columns don't show anything new, the right two

columns deal wit "assumptions" . This actually is an interesting approach as it covers aspects of a project which usually are tacitly, often unconsciously assumed (examples : HIV projects assume that victims care for their own health disregarding other goals, development plans assume that politicians, administrators care for the common goods etc.)

On a historical retrospective it is important to remark, that GTZ (one predecessor of GIZ) incorporated the logical framework or logframe approach into ZOPP. The original logframe had 16 cells containing the major elements of the management-by-objectives approach to project implementation. The matrix cells are organised in four

columns along a logical structure. The binding character for project planning of ZOPP has loosened and today "ZOPP must become more realistic and better account for social contexts".[10] advocating a pragmatic approach and a tacit denial that one planning, monitoring and evaluation system can or should be applied under all circumstances.

4.3. Systemic Approach by GIZ

"Development programs promote complex reforms and change processes. Today, such processes are characterized more than ever by insecurity and unpredictability, posing a big challenge to the evaluation of development projects. In order to understand which projects work, why and under which conditions, evaluations need to embrace the interaction of various influencing factors and the multi-dimensionality of societal change. However, present evaluation approaches often premise predictability and linearity of event chains. They reflect the natural human need for security but are often not suitable to comprehend complex situations.

In order to fill this gap systemic approaches in evaluation of development programs are increasingly being discussed. A key concept is interdependency instead of linear cause-effect-relations. Systemic evaluations look at interrelations instead of analyzing isolated facts

Module 2: Monitoring: international concepts and best practice using information

and figures. They focus on the interaction between various stakeholders with different

motivations, interests, perceptions, and perspectives.

4.3.1. Quality, results and evaluation at GTZ

Quality, results and evaluation are among the cornerstones of our work. They are closely

related and are important criteria for the success and substantiation of our services.

The quality of a product or service is measured by the extent to which it meets the

requirements of our clients and partners. We aim to provide services of an excellent quality

and to make continuous improvements to core and support processes. Our quality

management system is based on the internationally recognised model of the European

Foundation for Quality Management (EFQM). It comprises various measures designed on

the one hand to fulfil the quality requirements agreed with clients, and on the other to attain

the level of quality that is anchored in our rules and regulations. By these means, we ensure

a high degree of sustainability, cost-effectiveness, efficiency and effectiveness in our work.

Targeted steering of projects and programmes

Capacity WORKS is our management model for sustainable development at project and

programme level. It makes possible the effective and efficient steering of complex projects

and programmes, and it integrates our capacity development approach. Capacity WORKS

fixes the objectives and results agreed with the partner as key reference points, and then

provides management tools designed to achieve them. This objectives-oriented procedure

makes it easier to steer projects and programmes in terms of their quality and results.

Substantiating results

All our activities are intended to contribute to sustainable development around the world. The

effectiveness of our work is therefore a key requirement of our actions, which we use to

measure our success. In order to substantiate results on a systematic basis, we regularly observe and collect facts, figures and data and use them to indicate individual, institutional and systemic developments.

Conducting independent evaluations

In order to reliably, credibly and independently substantiate the results and successes of our work, projects and programmes are systematically examined in regular evaluations. International and local independent evaluators examine the results of our projects and programmes, while they are running (progress reviews), upon their completion (final evaluations) and several years after their completion (ex-post evaluations). Our evaluation system includes self-managed, decentralised reviews and evaluations that are undertaken by independent institutes.

By continuously collecting data on project/programme results (monitoring), periodically reviewing and assessing (evaluating) them and making them transparent, we substantiate the effectiveness of our work on a systematic basis. In this way, we account to our clients and the public about the expedient use of taxpayers' money. The monitoring and evaluation results also provide important lessons for improving the steering of projects and programmes and of processes throughout the company.

This interplay of quality management, Capacity WORKS, evaluation and results monitoring enables us to enhance the quality of our work durably and purposefully and thus to improve living conditions in our partner countries in an effective and sustainable way.

The System splits into three parts

4.3.2. Quality management at GTZ

What is quality management?

Defining features of GTZ quality management

Quality is defined in GTZ's Quality House model, which applies across the board to all business sectors. This model illustrates GTZ's structural features and is based on the concept of sustainable development. The model offers an internal framework of reference for monitoring and enhancing quality. It fits the criteria given in the EFQM Model of Excellence (European Foundation for Quality Management) and the success factors presented in Capacity WORKS, GTZ's own management model. GTZ thus implements quality management in line with the EFQM standard.

GTZ Quality Criteria

1.Values geared towards sustainability

2.Resources used economically

3.Efficient steering

4.Rules met

5.Positive results

GTZ Quality House
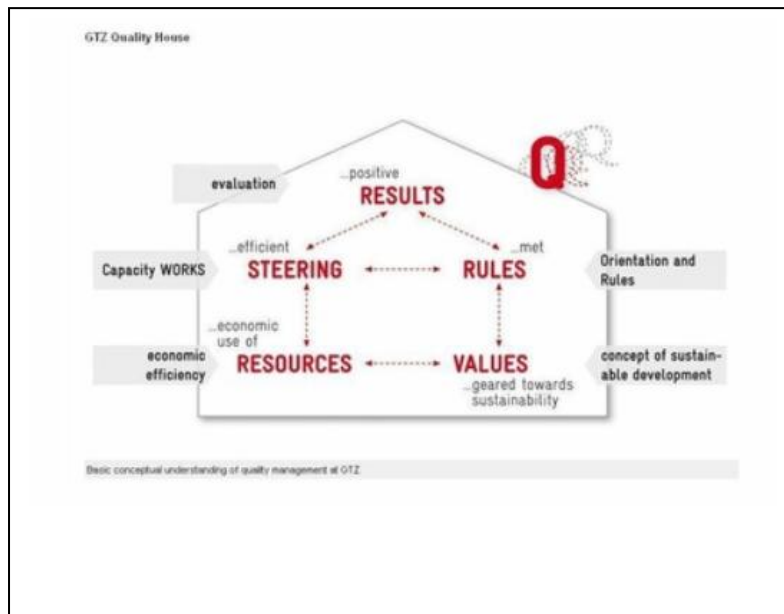
Instruments used in GTZ quality management

Quality report

The quality report gives GTZ an overview of the quality status of key service processes. Using concrete examples, it describes positive developments and future challenges for GTZ. The report is structured according to the five quality criteria.

Quality management, results-oriented monitoring, evaluation and Capacity WORKS

'Quality is when the customer comes back and not the goods.' True to this motto, quality at GTZ is defined as follows: all services are managed efficiently, delivered in an economical way, are results oriented and sustainable. This captures the essence of quality management at GTZ. It also ties in closely with results-oriented monitoring, which we use continually to obtain data on the results of our projects and programmes. To enable us to provide reliable, credible and independent proof of the results and achievements of our work, we regularly evaluate our projects and programmes.

The results and objectives of a programme are the central points of reference for Capacity WORKS, GTZ's own management model for sustainable development at project and

Module 2: Monitoring: international concepts and best practice using information

programme level. Capacity WORKS facilitates the effective and efficient steering of complex

projects, while integrating GTZ's capacity development approach.

4.3.3. Capacity WORKS -The Management Model for Sustainable Development

Capacity WORKS is GTZ's management model for sustainable development and helps

provide answers to the following questions: How does GTZ make an effective contribution to

its partners' capacity development? How do we structure the management and steering of

projects and programmes so that the results of capacity development are as effective and

sustainable as possible? Following a successful eighteen-month pilot phase, we have been

applying Capacity WORKS in the management of GTZ projects and programmes both in

Germany and abroad since January 2009.

Capacity WORKS is a response to changes in conditions in the international setting.

Structures both in partner countries and within the donor landscape have become more

complex. Normally a project or programme is now no longer designed and implemented with

just one partner, but involves a multi-organisational and multi-institutional framework.

GTZ's many years of experience in steering such multi-organisational contexts have

therefore been systematically analysed, providing the foundation for the model. The structure

of Capacity WORKS translates the corporate policy concept of 'sustainable development' into

practical questions.

The key points of reference for Capacity WORKS are the objectives and results jointly

agreed on in relation to the contract. In development cooperation projects, the procedure can

be continuously reviewed and corrected in relation to the five success factors presented in

the model. In each success factor, thoughts and actions are guided by key questions. This is

further supported by the GTZ-specific consultancy principles. Capacity WORKS opens up

various structured approaches that partners can apply flexibly in the management of complex

Module 2: Monitoring: international concepts and best practice using information

projects and programmes. The model enables users to identify, focus on and address the

relevant negotiation processes in these complex cooperation frameworks. Our partners are

empowered to develop their individual organisational capacities to the full.

Since 2008, Capacity WORKS has been a protected wordmark both in Germany and

internationally

4.3.4. Evaluation- Working on sustainable results: Evaluation at GTZ

Hardly any other area of policy involves so much evaluation as development cooperation. At

GTZ, too, systematically reviewing the success of our operations is an integral part of our

work. This is how we account to our clients and the public generally for the use of tax money.

Our clients, partners and the target groups in the partner countries place great trust in us. By

continuously examining the results and success of our work and making these transparent,

we aim to fulfil this responsibility and be credible partners in international cooperation.

Evaluation helps us to improve our work continuously, providing important information for

managing projects and programmes and promoting learning by individuals and the company

as a whole.

We use two different instruments for evaluating projects and programmes: decentralised

evaluations and independent evaluations. Partner institutions and local evaluators are

regularly included in these evaluations. Joint evaluations conducted with German and

international cooperation partners are becoming increasingly important.

Evaluations examine the following questions:

Are we doing the right thing? (relevance).

Are we achieving the project and programme objectives? (effectiveness).

Are we contributing to the achievement of overarching development objectives and results? (impact).

Is our action cost-effective? (efficiency).

Are the results durable? (sustainability).

GTZ's evaluation system follows the internationally recognised criteria and principles of the OECD Development Assistance Committee. In accordance with the principle of independence, the GTZ Evaluation Unit is clearly separated from the operative departments, and reports directly to the Office of the Managing Directors.

The term external evaluations is used for measures conducted by the German Federal Ministry for Economic Cooperation and Development (BMZ) within its Central Evaluation Programme. These measures are not covered by GTZ's evaluation system, and the quality assurance of projects and programmes is handled by independent auditors.

Evaluation System
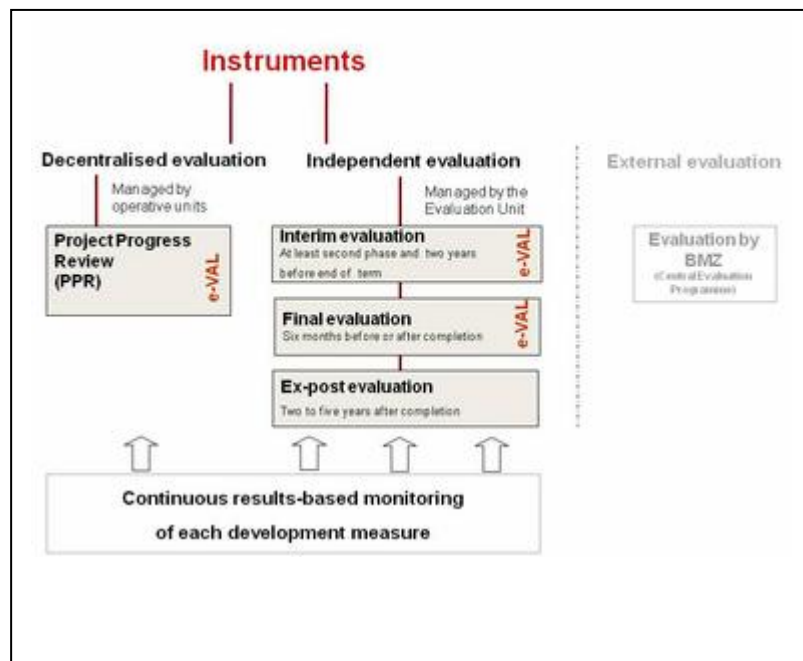
Assuring the Quality of Monitoring and Evaluation; Design and Methodology

This unit is about effective quality assurance (QA) of the monitoring and evaluation process and their related products. There is no reason to limit QA to the evaluation phase; likewise the quality of monitoring has to be assured. But for the sake of catching the important parts of the process we will stick to the proposals of UNDP.

5.1. *Monitoring for Results*

Here [11] provides step-by-step guidance on how to implement planned monitoring activities. It also presents useful tools and tips for effective monitoring and use of monitoring evidence in decision making.

Implementation of monitoring:

1. Have a clear common understanding of the following:

a. The monitoring policies applicable to the respective monitoring entity

b. Relevant roles and responsibilities and how they are applied in monitoring for both outcomes and outputs, and management entities in projects and programmes

c. Commonly used monitoring tools and approaches

2. Reinforce and elaborate the initial monitoring framework (as described before) with detailed information needed to implement monitoring actions. This includes finalizing reference points for periodic monitoring such as indicators, baselines, risks, and annual targets, and locking them in monitoring information systems.

3. Implement monitoring actions: organize, plan and implement monitoring actions, using selected tools for collection and analysis of data and reporting.

4. Use monitoring data objectively for management action and decision making.

Very important: There is no blueprint for monitoring that can be applied to all monitoring situations. The monitoring approach an organization uses in a given situation can be very different but should follow the general guidelines

5.2. *Evaluating for results*

Evaluation is critical to progress towards advancing human development.

Through the generation of 'evidence' and objective information, evaluations enable managers to make informed decisions and plan strategically. The success of any development program depends, in part, on the ability to carry out credible evaluations and use them to make evidenced-based decisions. The effective conduct and use of evaluation requires adequate human and financial resources, sound understanding of evaluation and most importantly, a culture of results-orientation, learning, inquiry and evidence-based decision making.

Norms for evaluation

Evaluation in UNDP should be:

- **Independent**—Management must not impose restrictions on the scope, content, comments and recommendations of evaluation reports. Evaluators must be free of conflict of interest
- **Intentional**—The rationale for an evaluation and the decisions to be based on it should be clear from the outset.
- **Transparent**—Meaningful consultation with stakeholders is essential for the credibility and utility of the evaluation.
- **Ethical**—Evaluation should not reflect personal or sectoral interests. Evaluators must have professional integrity, respect the rights of institutions and individuals to provide information in confidence, and be sensitive to the beliefs and customs of local social and cultural environments.
- **Impartial**—Removing bias and maximizing objectivity are critical for the credibility of the evaluation and its contribution to knowledge.
- **Of high quality**—All evaluations should meet minimum quality standards defined by the Evaluation Office
- **Timely**—Evaluations must be designed and completed in a timely fashion so as to ensure the usefulness of the findings and recommendations
- **Used**—Evaluation is a management discipline that seeks to provide information to be used for evidence-based decision making. To enhance the usefulness of the findings and recommendations, key stakeholders should be engaged in various ways in the conduct of the evaluation.

These principles are rather general but detailed examples also on various types of evaluation which will be discussed in another module are found in [11].

UNDP emphasizes the centrality of national ownership in evaluating results. The achievement of the outcome is dependent upon contributions from a range of partner. To this effect, the involvement of stakeholders and partners in the planning, management, conduct and use of evaluation is critical. The degree and modalities of their involvement will vary at different stages of the process. Some need only be informed of the process, while it would be important for others to be involved in a decision-making capacity. In each evaluation, a thorough assessment should be done in order to determine who the stakeholders are and

how they should be involved in the evaluation process. Again we will come back in detail with specific proposals for the evaluation process.

Proposed management structure for an evaluation

```
        ┌──────────────────┐         ┌──────────────────┐
        │ Program Owner like│         │ Program Partner  │
        │      UNDP         │         │  Co-Commissioner │
        └────────┬─────────┘         └────────┬─────────┘
                 └──────────────┬─────────────┘
                                │
  ┌────────────┐        ┌──────────────┐        ┌────────────┐
  │ Evaluation │────────│  Evaluation  │────────│ Reference  │
  │    Team    │        │   Manager    │        │   Group    │
  └────────────┘        └──────┬───────┘        └────────────┘
                               │
                       ┌───────────────┐
                       │Quality Assurance│
                       └───────────────┘
```
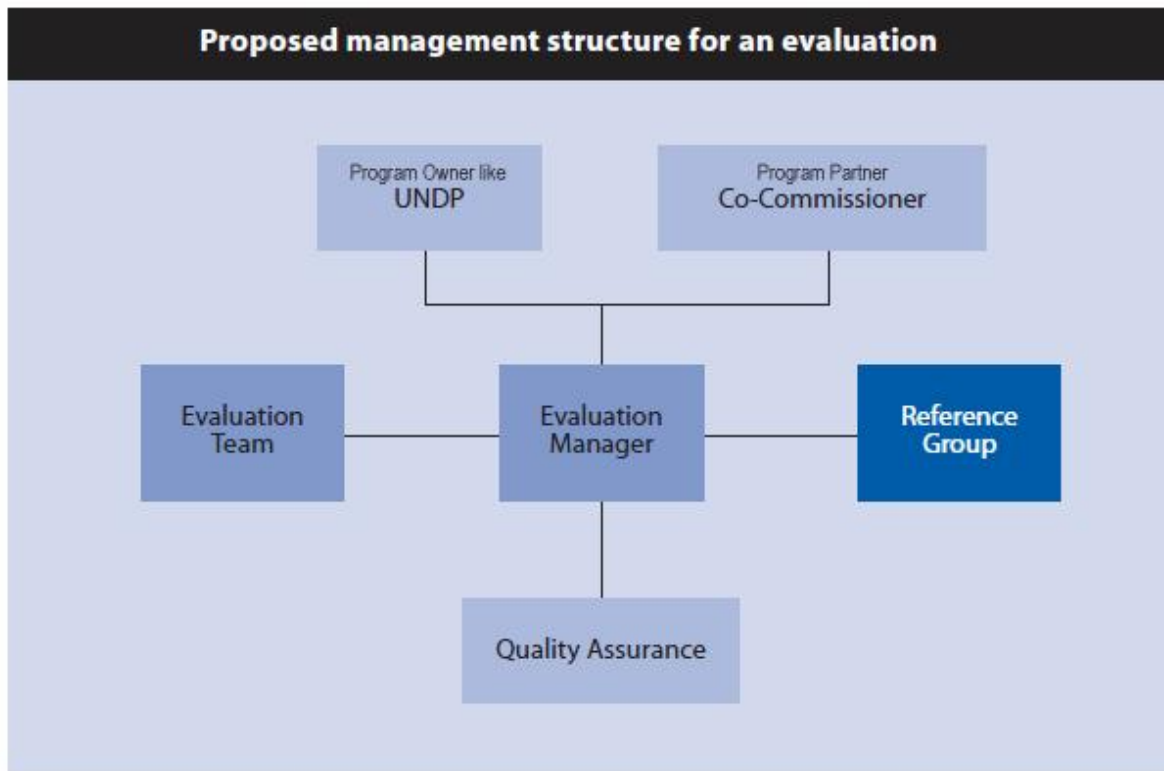
National ownership means that key partners and stakeholders must play an integral part in the evaluation process from the outset. For every evaluation, there should be a reference group comprised of key stakeholders to work closely with the evaluation manager to guide the process.

Developing a quality evaluation design involves a thorough understanding of what is being evaluated (the initiative and its context) and making decisions about the following key elements and how each will contribute to valid and useful evaluation results:

The purpose of the evaluation

The focus of the evaluation, that is, the key questions that the evaluation seeks to answer

The sources and methods for obtaining information that is credible and defensible

The procedures that will be used to analyse and interpret data and report results

The standards that must be reached for the initiative to be considered successful

The evidence that will be used to indicate how the initiative has performed and demonstrate its results (outputs and outcomes)

5.2.1. The Evaluation Purpose

All evaluations start with a purpose, which sets the direction. The evaluation risks being aimless and lacking credibility and usefulness without a clear and complete statement of purpose. Evaluations may fill a number of different needs. The statements of purpose should make clear the following:

Why the evaluation is being conducted and at that particular point in time

Who will use the information

What information is needed

How the information will be used

5.2.2. Evaluation Scope

The evaluation scope narrows the focus of the evaluation by setting the boundaries for what the evaluation will and will not cover in meeting the evaluation purpose.

5.2.3. Evaluation objectives and criteria

Evaluation objectives are statements about what the evaluation will do to fulfil the purpose of the evaluation. Evaluation criteria help focus evaluation objectives by defining the standards against which the initiative will be assessed:

Relevance

Effectiveness

Efficiency

Sustainability

Impact

5.2.4. Evaluation Methodology

The evaluation design must detail a step-by-step plan of work that specifies the methods the evaluation will use to collect the information needed to address the evaluation criteria and answer the evaluation questions, analyse the data, interpret the findings and report the results. There are many detailed issues in this chapter on data quality and how to assure it. This will be dealt with in detailed later in another module.

6. Conclusion

Having gone through four international Systems proposing on Monitoring and / or Evaluation strategies we can safely conclude that all cook according to a very similar recipe. We might exclude the rather exotic systemic approach of GIZ about which they state it is not considered a fit-for-all methodology. However we stated in this unit, that there is no one for all method and we have to adapt the general approach to the specific circumstances. We will try to apply the general thinking, and there is a lot of common thinking in all the international approaches to a country or region related exercise. The UN handbook not only provides a very comprehensive instruction, it is very up to date, it also offers a lot of templates to be used for control and selection of appropriate tools. Even though there is a chapter on Quality assurance, there is no comparable guidelines in this chapter to establish a QA standard. We will try to improve this in the appropriate module on evaluation and enhance it by proposing precise QA measure to assure the quality of both evaluation and monitoring.

Exercises and Group Work 1

Working groups to develop Timetable, Milestones and Quality assurance for M&E of a Development Subsystem in <country / region>, choosing one of the proposed Development Subsystems or for another of own choice:

- "Human development first", MDG

- Financial system, Banking System

- Economical Growth & Trade

- Regional development

Presentation of working groups results

Exercises and Group Work 2

Group Proposals for a Draft Concept of a Comprehensive Monitoring and/or Evaluation System for Development Planning of <country / region> relating to proposals and best practices of this module.

Presentation of working groups results

Bibliography

Documents and Links

The Prince (Italian: Il Principe) by Niccolò Machiavelli.

OECD Joint Evaluation of General Budget Support – 2006

http://www.oecd.org/document/33/0,3746,en_21571361_34047972_44005921_1_1_1_1,00.html

Evaluation of General Budget Support - Methods (OECD).pdf

Causality Map for the Enhanced Evaluation Framework.pdf

Joseph E. Stiglitz, Globalization and Its Discontents , W.W. Norton & Company

How to Build M&E Systems to Support Better Government IEG and WB,Keith McKay:

BetterM&Efor BetterGov (WB).pdf

WB_ResultChain for Building Govenment M&E.pdf

Monitoring & Evaluation, Some Tools, Methos and Approaches , WB :

M&E_tools_methods_approaches (WB).pdf

Ten Steps to a Results- Based Monitoring and Evaluation System, Jody Zall Kusek, Ray C.

Rist, WB: TenStepsToRB_M&E(WB).pdf

ZOPP Objectives-oriented project planning.pdf

HANDBOOK ON PLANNING, MONITORING AND EVALUATING FOR DEVELOPMENT

RESULTS United Nations Development Programme :   UN-handbook M&E.pdf

UNDP_Ex_MonitoringPlan.pdf

UNDP_Ex_EvaluationPlan.pdf

http://www.gtz.de/en/unternehmen/31523.htm

Better Monitoring for Better Evaluation

Module 3: Logical analysis - a way to improve monitoring

Contents

1. Introduction and aims

The aim of this module is to demonstrate that development programmes and projects benefit from having a clearly stated logical structure. We shall show that this can make it easier to put in place accurate and economic monitoring, although the model or the framework itself is only as good and effective as the people working with it. It will become clear that this way of reasoning is particularly useful when analysing the linkages and dependencies between the different parts of a development project or programme.

As is stated in the UNDP Handbook on Monitoring and Evaluation (2009):

- Without proper planning and clear articulation of intended results, it is not clear what should be monitored and how; hence monitoring cannot be done well.

We are not, in this module, necessarily promoting a formal Logical Framework Approach when designing projects and programmes. LF and LFA have many supporters and some critics and a great deal has been written already on the subject [see a selected reading list at the end of this module].

What we shall argue is that clarity of goal and purpose is essential for a development programme; and that the logical framework approach is a good way to bring this about. One only has to take the example of the Millennium Development Goal 7 target 7D

By 2020, to have achieved a significant improvement in the lives of at least 100 million slum dwellers

to see how necessary this is, and to understand why there is still discussion going on about what precisely the international development community is trying to do for slum dwellers. The key phrase is "a significant improvement in the lives". What do we want to improve? What level of improvement will be considered as significant? Is it permissible to improve different things for different groups of dwellers, and if so, how can we sensibly aggregate the results?

However, it is undeniable that a Logical Framework is a useful tool to help people formulate their ideas more clearly, and we shall go through this in more detail in Unit 2. In Unit 3, we shall examine the logical framework template now used by DFID UK, with an illustrated example. A number of key messages will be discussed in Unit 4.

Since accurate wording is the key to using a logframe correctly, Unit 5 will study in more detail examples of Goal, Purpose and Output statements from a number of logframes in use. Finally in Unit 6 we shall briefly discuss the other tools which are a part of logical analysis.

2. Logical Frameworks – history and development

2.1. Terminology

It is necessary to be clear about our terminology. LFA describes the Logical Framework Approach, which is an analytical process, and is hence sometimes called Logical Framework Analysis. One of its core components is the Logical Framework Matrix (LFM), which is often loosely called the Logical Framework, or, more familiarly, the logframe.

Our main interest in this module lies in the use of the logframe as a reality check, as an aid to clarity of objectives, and as an aid to formulating indicators. We shall begin, however, with a brief overview of the LFA. The key documents we have referred to are:

- The Logical Framework Approach (LFA): Handbook for Objectives-Oriented Planning, Fourth Edition, NORAD 1999

- The LogFrame Handbook: a Logical Framework Approach to Project Cycle Management, World Bank (undated)

- Aid Delivery Methods: Volume 1: Project Cycle Management Guidelines, European Commission March 2004

- AusGuidelines: Activity Design: The Logical Framework Approach, AusAid October 2005

- Guidance on using the revised Logical Framework, DFID February 2009

2.2. The early years

LFA was developed in the late 1960s by USAID, who wished to improve their project planning and evaluation systems. USAID were concerned that

1. planning was too vague, without clearly defined objectives that could be used to monitor and evaluate the success or failure of a project;
2. management responsibilities were unclear; and
3. evaluation was often an adversarial process, since there was no common agreement as to what the project was really trying to achieve.

LFA was enthusiastically adopted by most bilateral and multilateral development agencies, who added their own interpretations and levels of complexity.

LFA is both an analytical process and a set of tools used to support project planning and management. It is an aid to thinking, which allows information to be analysed and organised in a structured way.

2.3. Advantages of using LFA

NORAD highlights a number of advantages of using LFA, with which other agencies agree.

- It ensures that fundamental questions are asked and weaknesses are analysed, in order to provide decision makers with better and more relevant information

- It guides systematic and logical analysis of the inter-related key elements which constitute a well designed project

- It improves planning by highlighting linkages between project elements and external factors

- It provides a better basis for systematic monitoring and analysis of the effects of projects

- It facilitates common understanding and better communication between decision makers, managers and other parties involved in the project

- Management and administrations benefit from standardised procedures for collecting and assessing information

- The use of LFA and systematic monitoring ensures continuity of approach when original project staff are replaced

- As more institutions adopt the LFA concept it may facilitate communication between governments and donor agencies

- Widespread use of the LFA format makes it easier to undertake both sectoral studies and comparative studies in general

2.4.   ... but there are limitations too

The limitations of LFA, according to NORAD, are that:

- rigidity in project administration may arise when objectives and external factors specified at the outset are over-emphasised. This can be avoided by regular project reviews where the key elements can be re-evaluated and adjusted

- LFA is a general analytical tool. It is policy neutral on such questions as income distribution, employment opportunities, access to resources, local participation, cost and feasibility of strategies and technology, or effects on the environment

- LFA is therefore only one of several tools to be used during project preparation, implementation and evaluation ,and it does not replace target groups analysis, cost-benefit analysis, time planning, impact analysis, etc

- the full benefits of utilising LFA can be achieved only through systematic training of all parties involved and methodological follow-up

Source: NORAD 1999

2.5.   Comments from other agencies

AusAid adds that LFA establishes a logical hierarchy of means by which objectives will be reached. This hierarchy is an important point which we shall return to in Unit 4. They also point out that LFA is best started early in activity design, noting that it is difficult to use the LFA to review and restructure ongoing activities which were not designed using LFA principles and practices.

The European Commission stress the difference between the LFA, which is an analytical *process*, involving stakeholder analysis, problem analysis, objective setting and strategy selection, and the logframe itself, which, while requiring further analysis of objectives, how they will be achieved and the potential risks, also provides the documented *product* of the analysis.

The Logical Framework Approach	
Analysis Phase	Planning Phase
Stakeholder analysis – identifying and characterising potential major stakeholders; assessing their capacity	Developing logical framework matrix – defining project structure, testing its internal logic and risks, formulating measurable indicators of success
Problem analysis – identifying key problems, constraints and opportunities; determining cause and effect relationships	Activity scheduling – determining the sequence and dependency of activities; estimating their duration, and assigning responsibility
Objective analysis – developing solutions from the identified problems; identifying means to end relationships	Resource scheduling – from the activity schedule, developing input schedules and a budget
Strategy analysis – identifying different strategies to achieve solutions; selecting most appropriate strategy	

Source: European Commission, 2004

We shall return to these various analyses in Unit 6.

2.6. Evolution of logframes

In the early days of logframes, they were a simple 3x4 or 4x4 matrix, typically with the first column containing the narrative, the second for the indicators and means of verification (these were often split into two columns) and the final column used for the risks and assumptions.

An early logframe template

Narrative causality	Indicators and	Assumptions and risks

(mention target groups)	means of verification	
1. GOAL The higher-level objective towards which the project is expected to contribute.	1. GOAL INDICATORS Measures (direct or indirect) which verify to what extent the goal is fulfilled. Means of verification should be specified.	1. ASSUMPTIONS Important events, conditions or decisions necessary for sustaining the objectives in the long run.
2. PURPOSE The effect which is expected to be achieved as a result of the project.	2. PURPOSE INDICATORS Measures (direct or indirect) which verify to what extent the purpose is fulfilled. Means of verification should be specified.	2. ASSUMPTIONS Important events, conditions or decisions outside the control of the project which must prevail for the development objective to be attained.
3. OUTPUTS The results that the project management should be able to guarantee.	3. OUTPUT INDICATORS Measures (direct or indirect) which verify to what extent the outputs are produced. Means of verification should be specified.	3. ASSUMPTIONS Important events, conditions or decisions outside the control of the project management, necessary for the achievement of the immediate objective.
4. ACTIVITIES The activities that have to be undertaken by the project in order to produce the outputs.	5. INPUTS Goods and services necessary to undertake the activities.	4. ASSUMPTIONS Important events, conditions or decisions outside the control of the project management, necessary for the production of the outputs.

2.7. Writing a logframe

When filling in the logframe matrix, planners were encouraged to plan down the first column, to ensure a logical thread, and then analyse the risks and assumptions moving up the final column. Where possible, if risks could be mitigated and assumptions made less uncertain, by modifying the project design, this was to be done.

The terminology in the matrix above is reasonably clear. The word Purpose is sometimes replaced by Outcome, but this can easily be confused with Output. It is normal practice to use Outcome to define the measurable effect of the project at Purpose level, as follows:

The Purpose of the project is to ....; the Outcome will be ....

Once the narrative had been established, then indicators of achievement were to be entered, together with means of verification – in other words, how will we be able to measure progress at the various levels and where will that information come from.

2.8. Tackle the job early, and do it thoroughly

It is true to say that many people found logframes a difficult and even threatening concept. Some still do. Often, completing the logframe was left until the last minute – it was viewed as a "compulsory annex" to the project memorandum and not as a useful planning tool. The key thing is to examine the narrative causality to make sure that the steps from inputs to activities to outputs to purpose to goal make sense.

And because of the challenge to formulate sensible statements for goal and purpose – particularly for process projects or projects involving influence and institutional development – it was often the case that more emphasis was placed on inputs and activities – the easier part, but in fact much less important.

Similarly, more attention was usually given to the narrative and the statements of risk, and much less to the selection of suitable indicators and how to measure them – hence the need for this course.

3. The DFID logframe

3.1. Overview

DFID have recently (2009) revised their LogFrame Guidance to put much more emphasis on indicators, baselines, milestones and targets, and the activities have been moved to an annex. We shall use the DFID logframe as model for this course, as it encapsulates a number of important points, as explained in the following pages.

One important development was to restrict logframes to a single goal and purpose. Earlier examples were full of examples such as "we aim to do x, y and z", which means that there is a lack of focus. The argument is that it is better to have a package of projects and programmes, each with its own clear logframe, than to try to wrap up everything into one all-embracing framework. But as we shall see, we can also create an umbrella logframe for all the elements of the package, which focuses itself on the policy framework needed to ensure that all the components succeed and complement each other.

3.2. Logic

It is important to check the logic of the logframe. This is:

IF we undertake the activities AND the assumptions hold true, THEN we will create the outputs

IF we deliver the outputs AND the assumptions hold true, THEN we will achieve the purpose

IF we achieve the purpose AND the assumptions hold true, THEN we will contribute to the goal

It is crucial that there are no gaps in the logic.

3.3. A single Goal

The Goal is not intended to be achieved solely by the project. This is a higher-level situation that the project will contribute towards achieving. Project staff should be aware of other efforts being made to achieve the goal, so they can make informed decisions about neglected areas, and sectors where DFID and its partners would have the greatest added value. This information is also important when establishing the hierarchy of objectives (ie is your project nested within a broader undertaking?)

DFID also needs to keep track of overall progress towards the desired outcome, monitoring steps forwards and being mindful of steps backward. The Goal should therefore be measureable – in particular a measure that gives a steer on where we need to place renewed emphasis.

Indicators at Goal level should be Impact measures

3.4. ... and a single Purpose

There can only be one Purpose for the project. The Purpose should identify what will change and who will benefit. For development projects, the Purpose should refer to how the project will contribute to reducing poverty/contribute to the Millennium Development Goals

Indicators at Purpose level should be Outcome measures

3.5. Outputs

Outputs are the specific, direct deliverables of the project. If the outputs are delivered, then the conditions are in place which are necessary to achieve the purpose. The DFID template allows for up to ten outputs, but best practice suggests that if there are more than six, these should be examined closely with a view to reducing the number.

3.6. Indicators

Best practice again suggests that there should be a maximum of three indicators for each Output, and two each or Goal and Purpose.

Each indicator requires a baseline, a target, milestones and the source of data should be stated. Indicators, Baselines and Targets should be disaggregated by sex, age, beneficiary groups, etc, wherever relevant.

3.7. Process projects

Defining what results are to be expected from more process-driven projects may be

challenging, but DFID still needs to be able to report what we are achieving by channelling

UK funds into supporting government reform or influencing. For example, policy engagement

is likely to be effective when:

- it promotes evidence-based policy making;
- non-state domestic stakeholders are involved in the policy cycle and process;
- it is targeted, sequenced and long term; and
- it is based on clear channels of communication.

3.8. DFID Logframe Template (2009)

PROJECT TITLE						
GOAL	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	
		Source				
	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	
		Source				

PURPOSE	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	Assumptions
		Source				
	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	
		Source				
INPUTS (£)	DFID (£)	Govt (£)	Other (£)	Total (£)	DFID Share (%)	
INPUTS (HR)	DFID (FTEs)					

OUTPUT 1	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	Assumptions
		Source				
	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	
		Source				

Impact Weighting	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	
		Source				RISK RATING
INPUTS (£)	DFID (£)	Govt (£)	Other (£)	Total (£)	DFID Share (%)	
INPUTS (HR)	DFID (FTEs)					

OUTPUT 2++	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	Assumptions
		Source				
	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	
		Source				
Impact Weighting	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	
		Source				RISK RATING
INPUTS (£)	DFID (£)	Govt (£)	Other (£)	Total (£)	DFID Share (%)	
INPUTS (HR)	DFID (FTEs)					

DFID Activities Log Template

OUTPUT 1	ACTIVITY 1.1	Milestone 1	Milestone 2	Milestone 3	Risks	Monitoring officer
	ACTIVITY 1.2	Milestone 1	Milestone 2	Milestone 3		Monitoring officer
	ACTIVITY 1.3	Milestone 1	Milestone 2	Milestone 3		Monitoring officer

OUTPUT 2++	ACTIVITY 2.1	Milestone 1	Milestone 2	Milestone 3	Risks	Monitoring officer
	ACTIVITY 2.2	Milestone 1	Milestone 2	Milestone 3		Monitoring officer
	ACTIVITY 2.3	Milestone 1	Milestone 2	Milestone 3		Monitoring officer

3.9. A DFID Logframe completed

PROJECT TITLE						
GOAL	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	
		Source				
	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	
		Source				

PURPOSE	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	Assumptions
		Source				
	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	
		Source				
INPUTS (£)	DFID (£)	Govt (£)	Other (£)	Total (£)	DFID Share (%)	
INPUTS (HR)	DFID (FTEs)					

OUTPUT 1	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	Assumptions
		Source				
	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	
		Source				

Impact Weighting	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	
		Source				RISK RATING
INPUTS (£)	DFID (£)	Govt (£)	Other (£)	Total (£)	DFID Share (%)	
INPUTS (HR)	DFID (FTEs)					

OUTPUT 2++	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	Assumptions
		Source				
	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	
		Source				
Impact Weighting	Indicator	Baseline + year	Milestone 1	Milestone 2	Target + year	
		Source				RISK RATING
INPUTS (£)	DFID (£)	Govt (£)	Other (£)	Total (£)	DFID Share (%)	
INPUTS (HR)	DFID (FTEs)					

A DFID Logframe completed (continued)

OUTPUT 1	ACTIVITY 1.1	Milestone 1	Milestone 2	Milestone 3	Risks	Monitoring officer
	ACTIVITY 1.2	Milestone 1	Milestone 2	Milestone 3		Monitoring officer
	ACTIVITY 1.3	Milestone 1	Milestone 2	Milestone 3		Monitoring officer

OUTPUT 2++	ACTIVITY 2.1	Milestone 1	Milestone 2	Milestone 3	Risks	Monitoring officer
	ACTIVITY 2.2	Milestone 1	Milestone 2	Milestone 3		Monitoring officer
	ACTIVITY 2.3	Milestone 1	Milestone 2	Milestone 3		Monitoring officer

A DFID Logframe completed (continued)

4. Logical frameworks – key messages

4.1. General remarks

In the preceding units we have seen how logical frameworks have been developed by donor agencies into something which is intended to assist with the management of a development project or programme. It is worth remembering, though, that a logframe is not only for the use of development agencies and their programmes. Correctly used, a logframe (or its equivalent) is an aid to planning around which the project or programme team can manage, monitor, assess and ultimately evaluate progress.

Despite definitions of Goal, Purpose and Outputs which are apparently simple and clear, applying them in practice is not so easy. There is definitely a knack to formulating a purpose statement so that it is not just a list of things which are to be produced – these are the outputs. And it is equally hard to formulate an output statement so that it is not just a list of activities. We shall look at examples in Unit 5, and merely note here that great care should be taken with the words.

4.2. Ownership

It is interesting to reflect why logframes can be so unpopular among development practitioners, especially those who are close to the project or programme in question. There seems to be a feeling somewhere between dislike and fear which it is hard to pin down.

One possible answer is that the participatory nature of the logical analysis is being overlooked. As AusAid say:

As LFA is an 'aid to thinking', it has widespread and flexible application. Activity planning and management should always be approached as a team task. This means that adequate opportunity should be given to colleagues and key stakeholders to provide input to the process and product of LFA. This can be supported by

- Taking time to explain the principles of LFA and clarifying the terminology used
- Integrating effective team work and adult learning methods into meetings with stakeholder groups; and
- Ensuring that stakeholder groups are involved in situation and/or problem analysis, particularly in early design.

However, LFA is not a tool that all participants should necessarily be expected to understand or use. While 'logical' in concept, its effective application poses many challenges, even to the experienced user

It is too often the case that, while logical analysis is readily used at the design stage of a project or programme, the logframe itself is only completed at the end, as one of the many annexes of a project document (often, unfortunately, together with the proposed monitoring system). The participatory nature is then lost, and the logframe is felt to be imposed, or external, and not seen to be useful.

4.3. Comprehensiveness

Those close to the project (who can see the obvious benefits of the course of action proposed and feel that any further justification is unnecessary) sometimes feel that their aim

and objectives are being to forced to fit into an unsympathetic format which cannot reflect the complexity of the project.

This is understandable, of course, but the logframe is not meant to encapsulate the entirety of the programme. As DFID say, when used properly, "the logframe

- brings together in one place a statement of all <u>key</u> [our emphasis] aspects of the project in a systematic, concise and coherent way; and
- provides a framework for monitoring and evaluation where planned and actual results can be compared."

Given, then, that the logframe will only show what is key, it is even more important to get everyone's agreement on what is included and what is not.

4.4. Positioning

Another area which causes difficulty to those preparing logframes is what we shall call *positioning*, or what the World Bank describe as *primacy of point of view*. In short, one person's goal can be another person's activity. This is best illustrated with a (fictitious) example.

The Poverty Reduction Strategy Paper for UK has as its overall goal a desire to measurably improve the living conditions of UK citizens. The government will work towards this goal which is shared with the private sector, the IMF and the European Commission. The specific purpose of the strategy, for which the government is accountable to parliament, is to increase income levels across all sectors of society. Government feel, after extensive research and consultation, that three sectors need attention: education, where society needs more technically able students and fewer academics; health, where the high costs of maintaining

elderly people in homes is an unacceptable drain on both private and public resources; and transport, where there is an inefficient use of roads.

From the viewpoint of the prime minister, the EU president, and the managing director of the IMF, one of the outputs of the strategy will be more technically qualified students. They have little interest in how this is achieved, and will monitor the results by an annual headcount of students and a measure of quality.

Turn now to the viewpoint of the minister of education. Delivering more technically qualified students is his goal, shared with the universities and technical colleges and the finance ministry, monitored by the independent standards authority. The outcome, for which he is personally responsible, is to make sure that the educational system will deliver the results. To do this he needs changes in education law, in funding arrangements for colleges, and indeed more colleges.

The legal advisor to the minister of education has another viewpoint, at a lower level still in the logical hierarchy. Increasing UK income levels is for him a super-goal at best, and may not even appear on his personal logframe. Increasing the output of technically qualified students is a goal he shares with others. He is personally responsible for delivering the new legislation – that is his purpose statement – and will require a number of different inputs to do this.

And so the logical hierarchy continues down the page.

Here is another illustration taken from the UNDP Handbook.

Policy (of the National Agricultural Research Council)	Programme (of the Research Stations)	Project (of the Research Team)
Goal: to contribute to the improved livelihood of hill farming families		

Purpose: increased agricultural production, productivity and incomes among hill farming families	*Goal:* to contribute to increasing agricultural production, productivity and incomes among hill farming households	
Result: the use of improved agricultural technologies increased among targeted farmers	*Purpose:* increased use of improved agricultural technologies by hill farmers (erg rice)	*Goal:* to contribute to increased use of recommended improved technologies
	Result: recommendations for targeted farmers provided and disseminated	*Purpose:* recommendations provided for improved technologies suitable for targeted farmers
		Results: technologies identified based on farmer priorities; technologies generated and adapted; technologies verified in farmers' fields

4.5. Over-Ambition

In the DFID revised logframe template above, they say that:

For development projects, the Purpose should refer to how the project will contribute to reducing poverty/contribute to the Millennium Development Goals.

Taken together with the statement that there can be only one purpose, this leads to a tendency to sometimes situate projects within the hierarchy at a higher level than is sensible. If you use an MDG in your purpose statement, and therefore presumably include an MDG indicator, you are essentially saying that if the MDG indicator changes significantly (for better or worse) this can be linked to your intervention and you accept (or claim) some responsibility for the change.

Stakeholders will want to consider carefully whether this is accurate or not. It is sometimes better to refer to the MDGs at super-goal level, and have a goal statement which is less ambitious.

4.6. The link to indicators

Following on from the previous argument, it can be seen that correct positioning of the project has a direct impact on the choice of indicators which will best monitor progress.

In Module 5 we shall look in detail at the Millennium Development Goal indicators, which are high level indicators of general well-being, focused around the poorest in world society. Their utility is mainly as headline indicators whose aim is to shock and shame the more privileged, and developing country governments, into doing something about what one international summit after another have agreed are basic human rights – education for all, basic health care, a minimum income, and so on.

But these indicators have become embedded in development planning to the extent that people feel obliged to include them in their logframes, whether or not it is sensible. As we argued above, the inclusion of an MDG indicator in your logframe implies that a) you think your project can have an impact on it, and b) if it changes (in whichever direction), then you will take some share of the responsibility for that change.

It is hard to see how MDG indicators are appropriate for any level of a project below that of Goal, perhaps unless they are specific to a social group or geographical area. More frequently, logframe developers are now introducing a super-goal as a way to include an MDG without it unbalancing the logframe logic.

4.7. Process projects

Opponents of the logframe, or those who are less in favour of them, often claim that they are only suited to classical development projects where the outcomes are clear and measurable

by classical statistical means. This is not the case. We would argue, in fact, that the more nebulous projects of an influencing nature are more in need of a logical framework and some precision than the usual development ones.

We do not claim that the task is easy. Thought is required to formulate goal and purpose statements which accurately reflect why the money is being spent and what it is hoped to achieve. And some degree of imagination is required to select indicators which will reflect progress. We shall see in Module 5 that while the majority of indicators are those which can be quantified, there is increasingly a place for qualitative indicators, measuring changes in attitudes and relative preferences.

5. A closer look at Goal, Purpose and Output statements

5.1. Introduction

As we have discussed in earlier, care needs to be taken when formulating the goal, purpose and output statements. This applies equally to traditional development projects and high level strategic programmes. To reiterate:

- The goal should be a high-level aim, shared with other projects and programmes, but not so high that it is disconnected from the project or programme in question;

- The purpose should be something which the project or programme can achieve and which will contribute towards the goal in a manner which can be demonstrated. The purpose statement should say <u>what will change</u> and <u>who will benefit;</u>

- The outputs are what the project or programme will produce, and which will achieve the purpose; and

- All statements should be clear and accurate.

We follow with seven examples from development projects and programmes around the world, with comments. We have removed certain names and acronyms which might identify them.

5.2. Example 1

Goal	Purpose	Outputs
The people in [state] have improved health and nutrition status	Increased use of quality health, nutrition and sanitation services by the poor	1. Improved access to priority health, nutrition and water and sanitation services in underserved areas 2. Public health management systems strengthened

		3. Positive health, nutrition and hygiene practices and health-seeking behaviour of communities improved
		4. Improved use of evidence in planning and delivery of equitable health, nutrition and water and sanitation services

What will change and who will benefit? The target beneficiaries are the poor, although it would seem unlikely that non-poor would be turned away from the health centres. And what will change? This is a case where the purpose statement cause some doubt as to whether the focus is on getting people merely to <u>use</u> the services more frequently, or in addition to get the authorities to also <u>provide</u> good quality services. The word <u>quality</u> needs to be defined. The outputs show a focus on access and behaviour – the implication being that if more people seek out the services that will be a good sign of success. But the word <u>priority</u> in Output 1 also needs defining. Output 4 does not sit well with the other three – it could possibly be regarded as an input to Output 2.

5.3. Example 2

Goal	Purpose	Outputs
Government MDG targets 1 and 2 on income poverty reduction and hunger achieved by 2015.	1 million people in targeted areas have lifted themselves out of extreme poverty by 2015	1. Proven approaches to improving the livelihoods of the extreme poor taken to scale.
		2. Innovative approaches to improve the livelihoods of the extreme poor tested, evaluated and successes ready for scaling up.
		3. Increasing consistency in the understanding, sharing and application of approaches to addressing extreme poverty.
		4. Policy and practice at local and national levels shows increasing recognition of the needs of the extreme poor.

This is an interesting case where using the MDG targets as a Goal obscures the real thrust of the project which is about knowledge and its application. Both the Goal and Purpose statements look more like targets. Alternative statements could be:

The purpose of the project is to ensure that all available knowledge about improving the livelihoods of the extreme poor is used as effectively, efficiently and as widely as possible.

The goal of the project is to make a measurable improvement in the levels of income poverty and malnutrition for people in targeted areas.

5.4. Example 3

Goal	Purpose	Outputs
To improve the quality of education in [country]	To increase nation wide the access to basic education for children in [country] and the capacity of Ministry of Education to effectively develop Education plans and budgets	1. 2,801,619 primary school going children have access to learning materials and textbooks at 1 to 1 ratio 2. 5300 SDCs (school development committees) revitalised and functional leading to increased parents participation in school management. 3. National education recovery plan developed including a prioritisation and costing plan

In this example we have a double purpose, which most logframe guidelines do not recommend. There is also a big leap of logic between the purpose and the goal. The outputs 1 and 2 are a mixture of activity and output. The project is providing learning materials and textbooks – this can be an activity – and developing a national education recovery plan, which is also an activity. The main purpose would seem to be the effective use of planning and management of education, both at the ministry level and in the school development committees. The goal, surely, is to get more children into basic education (not merely providing access to it) and the overall super-goal is to improve education quality.

5.5. Example 4

Goal	Purpose	Outputs
To improve health & well-being of vulnerable households	To meet emergency and recovery needs of conflict affected communities in	1. Improved sustainable access to safe, reliable water supply and to safe sanitation facilities

in [state]	[state]	2. Improved health and hygiene related awareness and behaviours at community level
		3. Improved access to nutritional support
		4. Increased access to education infrastructure
		5. Improved food security and tree coverage

This project is centred around a conflict-affected state, and in particular the life in refugee camps. The goal and purpose statements are logical, but the outputs are oddly worded and look more like targets. Output 1 might be better expressed as provision of a sustainable supply of safe water and sanitation which is not more than some specified distance from the community. Usage and maintenance of the supply and facilities can then be indicators of access and sustainability, and a measure of whether the emergency needs are being met.

5.6. Example 5

Goal	Purpose	Outputs
To improve maternal, newborn and child health (MNCH) in [region]	To improve effective access to MNCH (including RI (routine immunisation)) services in four [sub-regions]	1. Strengthened [sub-region] and LGA (local government authority) governance of PHC (primary health care) systems geared to RI and MNCH
		2. Improved human resource policies and practices for PHC
		3. Improved delivery of MNCH services (including RI) via the PHC system
		4. Operational research providing evidence for PHC stewardship, RI and MNCH policy and planning, service delivery, and effective demand creation
		5. Improved information generation with knowledge being used in policy and practice
		6. Increased demand for MNCH (including RI) services
		7. Improved capacity of Ministry level to enable [regional] MNCH (including RI) activities

What will change and who will benefit? The beneficiaries are mothers and children in the sub-regions targeted. What will change is not so clearly stated.

There is too much information in this logframe. It seems that there is a policy and practice aspect (outputs 1, 2 and 7), plus a delivery and demand aspect (outputs 3 and 6 – though

output 6 could just as easily be an indicator of success), plus an information and knowledge aspect (outputs 4 and 5). As a result the project's real focus is unclear.

A simpler purpose statement could refer to improving the management and use of MNCH systems and information at all levels. The goal would be to get more people using MNCH and RI services, and the super-goal could be an improvement in health outcomes.

Alternatively, the logframe could be recast as several lower level logframes with an overall unifying logframe at the policy and management level.

5.7. Example 6

Goal	Purpose	Outputs
Incidence of extreme poverty in [country] halved by 2015	Three million people* in poor rural areas lift themselves out of extreme poverty and achieve sustainable livelihoods (by end 2013) (* 599,300 households)	1. 270,300 STUP (specially targeted ultra-poor) households supported to successfully develop sustainable income generating activities. 2. 329,000 OTUP (other targeted ultra-poor) households supported to successfully develop sustainable income generating activities. 3. Essential preventive, basic curative and promotive health care services ensured for 3 million ultra-poor people and wider community. 4. Awareness and confidence of ultra-poor people to achieve their rights and entitlements enhanced/access social justice strengthened. 5, Social mobilisation, advocacy and communication campaigns to influence policy makers, service providers and the mass of the people, in favour of the ultra-poor successfully implemented. 6. Gender equality successfully promoted and mainstreamed throughout the programme 7. Effective monitoring, research and evaluation systems operational and results effectively disseminated. 8. Programme management systems established and operational.

This logframe has too much information. The key outputs would seem to the first three.

Everything else could be supporting activities.

5.8. Example 7

Goal	Purpose	Outputs
Sustainable improvement in the lives of all orphans and other vulnerable children (OVC) in [country]	Increased use of basic social and protective* services by OVC * Birth registration, child protection, education, basic health, psycho-social , food and nutrition, life skills and vocational training, child participation	1. Increased numbers of OVC provided with free social services and protection from abuse 2. Structures for effective national coordination and management of the programme are in place, and able to generate relevant policy and programme guidance for OVC interventions. 3. Strengthened capacity of programme partners supporting OVC 4. Programme wide M&E systems in place to measure impact and ensure best practice interventions

This logframe has a clear focus, and only output 1 sits uneasily with the others. It would be better used as an indicator for the purpose statement.

6. Other structural analysis tools

6.1. Introduction

In Unit 2 we introduced the logical framework and the part it plays in LFA, or logical framework analysis. We recall that the European Commission summarise the approach as follows:

The Logical Framework Approach	
Analysis Phase	Planning Phase
Stakeholder analysis – identifying and characterising potential major stakeholders; assessing their capacity	Developing logical framework matrix – defining project structure, testing its internal logic and risks, formulating measureable indicators of success
Problem analysis – identifying key problems, constraints and opportunities; determining cause and effect relationships	Activity scheduling – determining the sequence and dependency of activities; estimating their duration, and assigning responsibility
Objective analysis – developing solutions from the identified problems; identifying means to end relationships	Resource scheduling – from the activity schedule, developing input schedules and a budget
Strategy analysis – identifying different strategies to achieve solutions; selecting most appropriate strategy	

Source: European Commission, 2004

We shall now look briefly at each of these in turn. It is not our intention to go into great detail, as that can be easily found in the documentation. But for the purposes of this module it is useful to show the logic which should underpin a well-designed project or programme.

6.2. Stakeholder analysis

The stakeholders in a project or programme are the people and institutions who have an interest in the operation and outcome of a project, and include those who may lose out if it is successful – perhaps entrenched interest groups – as well as the expected beneficiaries. Their concerns and interests need to be explicitly understood and addressed in the project formulation, the aim being of course to maximise the benefits for those targeted and to minimise the potential negative impacts.

A key part of stakeholder analysis is gender analysis, which has its own tools and techniques, and with that general considerations of equity, ensuring that the benefits of the project reach vulnerable groups – often the poor, children, women and the disabled,

6.3. Problem analysis

Problem analysis identifies the negative aspects of an existing situation and tries to establish cause and effect relationships between the identified problems. The analysis is presented in the form of a Problem Tree, which sets the problem at the top and shows its causes below. The aim is to identify the real bottlenecks to which stakeholders attach high priority, and which they want to overcome. Problem analysis is only effective if carried out in a fully participatory manner.

As an example, let us say that the key problem which the group has identified is that not enough girls are attending school. Why? Among the reasons are: lack of toilet facilities for girls; lack of women teachers; a preference for spending limited resources on education for sons. Why are there no toilets for girls? Why are there not enough women teachers? Why is there a son preference? And so questions and answers go on down to the roots of the problem tree.

The aim is not to try to show in detail the whole situation, but to tease out the essentials of the existing negative situation. It is in fact the most crucial step of project formulation, as all subsequent analysis and decision-making rests on a correct statement of the problem and its root causes – a statement which is accepted by stakeholders.

6.4. Objective analysis

Once the problem tree has been agreed, the language in it is reversed from negative to positive to demonstrate the situation some time in the future when the identified problems have been remedied.

In the example above, lack of toilet facilities for girls can be restated as every school has toilet facilities for girls.

So as each problem is restated, it becomes an objective of the proposed project, and the logical connections can be verified again, to ensure that if one objective is achieved, then that is sufficient to ensure that the next one up the tree can be achieved. At this stage, realism can come into the discussion – for example, e*very school has at least one girls' toilet.*

6.5. Strategic analysis

The final step is to choose the best way to achieve the objectives, since a number of possibilities will have arisen during the various discussions. For instance, in our example, should the education authority build the toilets for girls or should the work be done by concerned parents? Does land need to be purchased? The questions which arise include:

- Should we try to solve all the problems, or just a selection?
- Are their positive opportunities which can be built on?
- What combination of interventions is most likely to bring about the desired results?
- How can local ownership of the project be supported?
- How can the capacity of local institutions be enhanced?
- What are the cost implications?

6.6. Results chain

DFID, in their recent logframe guidance note, refer to Results Chains. The Note says:

The logframe is an expression of the Results Chain – the results you expect the project to achieve. The box below shows how it aligns with the logframe format. You will see that the Purpose level in the logframe should be populated with Outcome level indicators, and the Goal level with Impact level indicators.

Results chain	Inputs	Process	Output	Outcome	Impact
Logframe	Inputs	Activities	Output	Purpose	Goal

The Results Chain must be based on evidence about what has worked in the past, so this is a real opportunity to take account of all the lessons learned, evaluation and research evidence available that underpins the design of the project. The evidence will also enable you to identify realistic targets: how much change does evidence suggest might be achieved over the project period?

Results chains are therefore a portion of the Strategic Analysis. They stress a linear chain of action which in most cases is too simplistic.

6.7. Discussion

The key message of what has been written earlier is that clarity of objectives, based on a shared understanding of the problem to be addressed, expressed in clear and accurate language, and presented within a logical structure, will assist immeasurably with the formulation of indicators to monitor progress.

There is no doubt this is easier to do for traditional development projects; but there is also no doubt that it is even more necessary to do for programmes which support national strategies, or which seek to influence or change behaviours. The key thing is to be honest about what you think you can achieve, and not to claim to be able to change a situation unless you are prepared to take responsibility for that change, whether good or bad.

6.8. Summary

We finish with a useful summary from AusAid.

	Differences in analytical focus	
	Macro-policy and programme support	Projects and stand-alone technical assistance
Problem Analysis and Stakeholder Analysis	Focus more on: Macro-economic framework, status of the aid 'market', national development/poverty reduction strategies, sector programme strategies, public finance management systems, institutional framework, organisational capacity, donor coordination, etc	Focus more on: Overview of institutional and organisational context, then with greater attention paid to specific technical constraints or problems within the focus area of 'intervention'
Objectives Analysis	Focus more on: Partner Government priorities, higher level policy and programme level objectives, development outcomes/results, strategies for achieving results	Focus more on: Project-based objectives, technical outputs, activities and input requirements
Strategy Options	Focus more on: Analysis of options for working through or within Partner Government systems, coordination arrangements with other development partners, aid effectiveness considerations, etc	Focus more on: Choosing between technical alternatives, considering appropriate parallel management and financing arrangements, considering Australian comparative advantage

Performance Indicators and Means of Verification	Focus more on: Higher level indicators of impact and outcome (results), use of MDGs or similar, using (or building on) existing Partner Government systems, harmonisation with other donors, accountability to 'beneficiaries'	Focus more on: Project specific indicators, output focus, contract milestones, reporting to Government of Australia, attribution of benefit to Australian contribution, accountability to Australian tax payers/authorities
Assumptions and Risks	Focus more on: Assumptions from the perspective of the partner government and their 'programme managers'	Focus more on: Assumptions from the perspective of the 'project managers' and the donor(s) - as well as on those from the perspective of the partner government and their 'programme managers'

Bibliography

The Logical Framework Approach (LFA): Handbook for Objectives-Oriented Planning, Fourth Edition, NORAD 1999

The LogFrame Handbook: a Logical Framework Approach to Project Cycle Management, World Bank 2004

Aid Delivery Methods: Volume 1: Project Cycle Management Guidelines, European Commission March 2004

AusGuidelines: Activity Design: The Logical Framework Approach, AusAid October 2005

Guidance on using the revised Logical Framework, DFID February 2009

Better Monitoring for Better Evaluation

Module 4:   Case studies (1): designing logframes to improve monitoring

Contents

1. Introduction, aims and methodology

1.1. Introduction and aims

The aim of this module is to give participants practical experience in developing a logframe based on two real life worked examples.

The key documents for this module are:

Republic of The Gambia Education Sector Medium Term Plan 2008-2011

Health Sector Reform Programme: Strategic Thrusts with a Logical Framework and a Plan of Action, 2004 – 2007, Federal Ministry of Health, Abuja

Both documents outline a relatively complex sectoral strategy, with an analysis of the problems the sector is facing and proposals for improvement. The Gambia Strategic Plan contains no logical framework at all, though it has a large number of activities and a list of monitoring indicators. Frequent references are made to wider aims and objectives.
The Nigerian document contains a number of logical frameworks for components of the overall strategy, though these do not quite follow standard logframe guidance. There is no overall logframe for the programme.

The exercise is to summarise each plan in the form of a logical framework, using the DFID model described in Module 3. We are not trying to describe the entirety of the project or programme – that is what the project documents are for.

At this stage, the monitoring indicators and data sources are not required: the key task is to formulate sensible Goal and Purpose statements and statements describing the main Outputs, preferably no more than four in number.

1.2. Suggested methodology

There are no hard and fast rules for doing this work. In an ideal world, the logframe is an integral part of the project document, and is developed by a meeting of stakeholders as part of an iterative process. Doing it in isolation from stakeholders is far from ideal, and is only made necessary by the nature of this training course. Government support is crucial and they are clearly among the stakeholders.

It is strongly recommended that wherever possible, the exercise is carried out by group working. If that is not possible, then an attempt must be made to share ideas with others.

As will be seen during the worked examples, there is no single correct answer to What is the Goal? and What is the Purpose? The statements have to be developed so that as many of the stakeholders as possible can buy into the final statement.

Having said that, here is one possible approach.

- Study the document carefully with a highlighter pen ready.

- Look for "we will do" statements, which are usually promises of activity, "we will achieve" statements, which are usually promises of output, "we aim to" statements, which are usually hopes for outcomes. Extract these into separate groups for ease of comparison.

- Look for statements containing phrases such as "our goal is to", "our purpose is to". Check carefully that these are correct usages of the words goal and purpose and amend if needed. Extract these.

- Be careful with statements which say "our objective is to". These can refer to outputs as well as outcomes (purpose), or for that matter to an overall goal.

- Make a note of statements referring to risks and assumptions.

- Bring all the statements together to look for consistency and coherency. Are some statements clearly more powerful than others? Are some statements clearly higher up the logical hierarchy than others?

- Is there a single high level statement which would encapsulate the purpose? If not can one be created?

- Can the outputs be grouped into three or four coherent sets, and an output statement formulated for each set? If not, can some of the outputs be described instead as activities which support more important outputs? Or simply not mentioned. [For instance, reforming the management skills of a ministry could be a desired outcome, but could equally well be an output to support the purpose; or simply an activity which underpins a number of outputs – it depends on your point of view].

- Is it necessary to develop a nested set of logframes, in cases where the project or programme is too complex to be summarised with just one logframe? In which case, formulate the goal and purpose of the highest level one first, then the outputs. Afterwards, let the purpose of the highest level logframe become the goal of each logframe at a lower level, and reformulate the outputs into a purpose statement.

- Test your understanding of the issues and the main thrust of the project or programme with a stakeholder group. Reformulate statements. But don't allow the logframe to be obscured with too much detail. In particular do a reality check on the purpose statement – it should be something which the programme managers feel they are responsible for achieving. If things go right they are happy to take the credit. If things go wrong they are prepared to take the blame.

- Check the logic of the risks and assumptions column to make sure nothing has been overlooked. Ask the stakeholder group to complete the logical statements: "if the described activities are done we shall achieve the output, assuming that ....". "If the outputs are produced then we shall achieve the purpose, assuming that ....".

2. Republic of The Gambia Education Sector Medium Term Plan 2008-2011 (MTP)

The original document is Reference A.

Reference B is the same document with highlighted points and commentary.

Reference C is the original Education Sector Strategic Plan 2006-2015 (ESSP), which the MTP complements.

Here is one possible line of reasoning towards creating a logframe, using the highlighted comments in Reference B.

2.1. Goal of the MTP

First we look for any statements in the text which could become an MTP Goal statement. There are several, as follows:

Reference comment PC2:

"the guiding philosophy for the development of the current education policy is *"Rethinking Education for Poverty Reduction".*"

> This refers to the ESSP, not the medium term strategy. But could this provide the basis of a goal or super-goal statement?

Reference comment PC5:

"[referring to VISION 2020]...sustained by a well-educated, trained, skilled [...] population,"

> Providing a well-educated, trained and skilled population could be a Goal or a Purpose for the ESSP.

Reference comment PC6:

"Thus, the ESSP is grounded in the overall goal of ensuring that 'by 2015 universal access to relevant and high quality education has been achieved.'

> ESSP goal or super-goal statement. Linked to MDG.
>
> Reference comment PC8:
>
> "education provision in The Gambia aims at reducing poverty."
>
> Another potential Goal or super-goal.

Reference comment PC9:

"the sector is guided by the following motto: "Provision of Responsive, Relevant and Quality Education for All Gambians for Poverty Reduction.""

This could almost be a goal statement. But is the stress on education or poverty reduction?

2.2. Discussion on a Goal statement

We recall from Module 3 that the Goal should be a high-level aim, shared with other projects and programmes, but not so high that it is disconnected from the project or programme in question.

The main decision to make is whether the MTP has the same Goal as the ESSP, or whether we situate the strategy a little lower in the logical hierarchy, or a little to the side. After all, the ESSP has not been dropped, and runs for a longer period than the MTP. Also, the ESSP

covers all aspects of education whereas the MTP has a focus on Basic education and one or two other areas of concern.

The ESSP does not have a logframe. If it did, it might have a super-goal which embraces the poverty reduction aspect (comments PC2, PC8 and PC9), and a Goal focusing on the educational aspect (comments PC5 and PC6), such as:

ESSP Super-Goal:

Reduction of poverty (through provision of responsive, relevant and high-quality education for all Gambians)

ESSP Goal:

A: to provide a well educated, trained and skilled population in support of VISION 2020; or

B: to achieve universal access to relevant and high-quality education by 2015.

The second statement is weaker than the first since it concentrates on access, ie the availability of education, rather than actual use of the facilities. As such, Goal A is stronger and more challenging. However, Goal B links to Millennium Development Goal 2, *Achieve Universal Primary Education*, whose related target is *Ensure that, by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling*, and this linkage can make it more attractive to donors.

Note that attaining either Goal is not within the sole gift of the ESSP, but is shared with donors, the private sector, the Finance Ministry, the Planning Ministry, the outcomes of the Gambia PRSP, and so on, which is why it is a Goal and not a Purpose.

Now we can turn to the MTP.

It is clear from the text that the focus is on Basic Education. We therefore have a choice. Either 1) we can regard quality basic education as the principal component of the education system; or 2) we can regard basic education as one of the pillars of the ESSP, along with Secondary and Tertiary education.

If we see it as a major component, then the ESSP and MTP share their super-goal and goal statements, with the MTP making specific reference to Basic education.

If we see it as one of the pillars, then the Goal of the ESSP could be the super-goal for the MTP; and the Purpose of the ESSP could be the Goal of the MTP.

This dichotomy is easiest shown in a table.

ESSP	MTP option 1	MTP option 2
Super-Goal: reduction of poverty (through provision of responsive, relevant and high-quality education for all Gambians)	Super-Goal: reduction of poverty (through provision of responsive, relevant and high-quality **basic** education for all Gambians)	
Goal A: to provide all Gambians with the good quality education, training and skills required by VISION 2020	Goal A: to provide all Gambians with the good quality **basic** education required by VISION 2020	Super-Goal A: to provide all Gambians with the good quality education, training and skills required by VISION 2020
Goal B: to achieve universal access to relevant and high-quality education by 2015.	Goal B: to achieve universal access to relevant and high-quality **basic** education by 2015.	Super-Goal B: to achieve universal access to relevant and high-quality education by 2015.
Purpose:	Purpose:	Goal: to be defined, but will contain a specific reference to basic education

At this stage we have to leave the table until we have considered potential Purpose statements.

2.3. Purpose of the MTP

We could first ask why it was found necessary to write the MTP. There is an Education Sector Strategic Plan 2006-2015 (ESSP) so why was this not sufficient? Can this point us to a Purpose statement?

In the Foreword by the Secretary of State, there is the statement that a <u>realistic</u> strategy is needed, implying that the ESSP was too ambitious. It goes on to say that some of the strategies in the ESSP have been undermined by recent developments. One example quoted later is the success of the Girl-Friendly School Initiative, which has led to a drop in enrolment by boys. The MTP is therefore a focusing and tightening of the ESSP, with greater realism. There is also a strong stress on what is defined as Basic education.

But this is not sufficient to say what the Purpose might be. One could say that the purpose of the MTP is to put something extra into the education system in The Gambia, in order to ensure that the ESSP achieves <u>its</u> purpose. An alternative view could be that the purpose of the MTP is to achieve something concrete for Basic education, thus contributing to the wider aims of the ESSP.

Are there any other statements in the text which could become an MTP Purpose statement? There are several, as follows:

Reference comment PC7:
"the guiding precept is the Mission Statement of the Department of State for Basic and Secondary Education (DOSBSE), which seeks to:
Provide access to relevant and high quality basic education for all
Provide high quality education services
Ensure gender equity in education

Provide relevant life skills

Promote the principle of lifelong learning."

> The mission statement could be a Purpose; but has too many components. It gives detail to the ESSP overall Goal [see comments PC5 and PC6 below], but does not add anything new, merely focusing on basic education.

Reference comment PC11:

"The provision of quality education is essentially the main thrust of the mandate of DOSBSE."

If quality is the main thrust should this be the purpose of the strategy; or is provision of

quality basic education an output to support a higher purpose?

Reference comment PC17:

"The Gambia remains highly committed to developing its human resource base with priority

given to free basic education for all."

Developing the Human Resource Base through free basic education could be a Purpose

statement

Reference comment PC21:

"Meeting Human Resource Needs Through Post-basic Education"

This seems to be a Purpose statement, but it is not obvious what it exactly means and would

be difficult to monitor without greater clarity.

Reference comment PC23:

"During the National Education Policy the GOTG will ensure that the higher education sub-sector provides a flexible and dynamic system of education and training that will address the demands of access and equity on the one hand, and the need for quality and standards (excellence), on the other."

This could be a Purpose statement for the Higher Education sector.

Reference comment PC30

"Gender Mainstreaming. Before the advent of the current education policy, the sector developed and implemented a comprehensive girls' education programme whose main objective was to bridge the gaps in enrolment, retention and performance between boys and girls in basic and secondary education. However greater prominence was given to basic education."

The purpose of the gender component of the ESSP.

Reference comment PC36:

"Such an ambitious vision needs to be matched with an education system that can ameliorate both the economic conditions and the skills of the populace."

A purpose statement for Science and Technology Education.

Reference comment PC37:

"The main objective of improving the management of the sector is to ensure effective and efficient delivery of education services."

This could be rephrased as an output supporting the higher level purpose of quality education outcomes.

2.4. Discussion on a Purpose statement

There are three types of reference above. References 7, 11 and 17 refer to the MTP as a whole. References 21, 23 and 36 refer to certain aspects of post-basic education. References 30 and 37 refer to cross-cutting themes of gender and management.

Turning first to Reference 7. This is the mission statement of the department with main responsibility for ensuring the MTP is a success. As a mission statement it is a bit lengthy, and there is repetition. It would seem unlikely that high quality education can be delivered without high quality education services; and education for all implies gender equity. Relevant basic education should of necessity contribute to relevant life skills, so the principle of life-long learning is the only point which is not encapsulated by the first bullet.

But the Purpose statement should provide the focus of the programme: and not try to cover every small detail. Therefore, rewriting the first bullet point slightly, we can start by suggesting:

(1) The Purpose of the MTP is to deliver relevant and high-quality basic education to all Gambians.

Looking now at references 11 and 17, are there any additional points which could be added to the purpose statement? The provision of quality basic education is implicit in delivery. Developing the human resource base is linked to the relevance of the education. The one addition is the word free.

Our Purpose statement could then be modified to read:

(2) The Purpose of the MTP is to deliver relevant and high-quality free basic education to all Gambians.

What of the other references? References 21, 23 and 36 cover post-basic education, higher education and science and technology education. These are no doubt important, but reading the MTP they take second place to basic education. Is it necessary then, for them to appear in the logframe, and if so, how?

If they are incorporated into the Purpose statement, it will become unwieldy. We would have a text something like:

(3) The Purpose of the MTP is to deliver relevant and high-quality free basic education for all Gambians, together with appropriate post-basic, higher and science and technology education.

This unwieldiness is due to the fact that the MTP is in fact a number of separate but linked programmes, being a package of remedial activities designed to correct the faults in the ESSP. Returning to an earlier remark, we could argue that the purpose of the MTP is to ensure that the ESSP is successful.

(4) The Purpose of the MTP is to provide necessary additional inputs to the education system in order that the ESSP will reach its stated Aims.

Finally, we look at references 30 and 37. These cover gender mainstreaming and management. In the logframe structure we are developing, these do not need to be included at Purpose level. It is clear that without good management the strategy will fail, and without

gender mainstreaming the equality aspect will not be achieved. These, then, can be regarded as outputs which support the purpose.

2.5. Summary so far

We have two possible candidates for the Purpose statement, (2) and (4) above. (4) wraps up all the components of the MTP and is flexible enough to allow for further issues to be addressed. But it is not very specific and would be extremely hard to monitor. (2) puts the emphasis on basic education which is the main remit of the ministry, but we risk losing sight of the other components if they are not included.

It is at this stage, or earlier, that consultation with the stakeholders, and in particular those charged with delivering the MTP, is needed to make a decision.

Let us now insert purpose (2) into a draft logframe and see how the logic looks. We shall use MTP Option 1 for the exercise.

Logframe (1)

Super-Goal	Reduction of poverty (through provision of responsive, relevant and high-quality **basic** education for all Gambians)
Goal	To provide all Gambians with the good quality **basic** education, training and skills required by VISION 2020
Purpose	To deliver relevant and high-quality free basic education to all Gambians

It will be seen that the goal and purpose statements are too similar. Delivery and Provision are too alike. If we accept the Purpose statement, then we need to modify the Goal to provide the stepping stone between delivering basic education and reducing poverty. A possible modification could refer to improvements to the human resource base, and another

would remove the reference to basic education, which makes the link between the MTP and the ESSP more explicit. Finally, we could make the super-goal statement simpler.

We then have:

Logframe (2)

Super-Goal	Reduction of poverty in The Gambia
Goal	To create a human resource base in The Gambia with the education, training and skills required by VISION 2020
Purpose	To deliver relevant and high-quality free basic education to all Gambians

It will be seen that this is one possible set of statements which meet the logframe criteria. The Purpose is specific to the MTP, and the ministry can be held responsible for its success, or otherwise. The Goal is shared with the ESSP, the wider education community, and others. The super-goal provides the link to the PRSP.

2.6. Outputs of the MTP

Turning now to Outputs, we recall from Module 1B that these are the results that the project management should be able to guarantee. If the outputs are produced, then the Purpose will be met.

What statements are in the text which could refer to outputs?

Reference comment PC10:

"In pursuance of the education sector's goals, aims and objectives, several key areas have been prioritised in the ESSP. These include access; quality education; teacher education,

Module 4 : Case studies (1) : designing logframes to improve monitoring

deployment and utilization; technical and vocational education and training; higher education

and sector management."

> Compare with Appendix 5: Access, Equity, Quality, Relevance, Efficiency, Resourcing, Management & Governance
>
> Reference comment PC11:
>
> "The provision of **quality education** is essentially the main thrust of the mandate of DOSBSE."
>
> If quality is the main thrust should this be the purpose of the strategy; or is provision of quality basic education an output to support a higher purpose?
>
> Reference comment PC14:

"In spite of the achievements made over the years, critical challenges still remain with the

sector. These embody matters of access, quality, resources and management."

Compare with the ESSP Priorities in comment 10

Reference comment PC16:

"Achieving Universal Completion of Quality Basic Education"

> This statement embodies access for all; equity; quality; and impact (through the reference to completion).
>
> Reference comment PC23:
>
> "During the National Education Policy the GOTG will ensure that the higher education sub-
>
> sector provides a flexible and dynamic system of education and training that will address
>
> the demands of access and equity on the one hand, and the need for quality and standards
>
> (excellence), on the other."
>
> This could be a Purpose statement for the Higher Education sector.
>
> Reference comment PC25:

"Under the basic education programme, the following cross-cutting interventions will be key in the realization of the provision of nine years of uninterrupted basic education of good quality:

- Quality
- Classroom construction
- Gender mainstreaming
- Education sector response to HIV/AIDS
- Life skills education
- Special needs education
- Child protection
- Teacher education
- Science and technology"

Many or all of these could be regarded as activities which support the higher basic education sector strategy.

Reference comment PC30:

"Before the advent of the current education policy, the sector developed and implemented a comprehensive girls' education programme whose main objective was to bridge the gaps in enrolment, retention and performance between boys and girls in basic and secondary education. However greater prominence was given to basic education."

The purpose of the Gender component.

Reference comment PC34:

"Reaping the full benefit of any country's youth will be dependent on, inter alia, the need to provide them with both livelihood and life skills. This will no doubt build individual capacities for meaningful participation in the national economy"

A desired output for Life Skills component.

Reference comment PC37:

"The main objective of improving the management of the sector is to ensure effective and efficient delivery of education services."

> This could be rephrased as an output supporting the higher level purpose of quality education outcomes.

2.7. Discussion on Output statements

The words and phrases which appear most often in the references above are:

1. Quality
2. Access
3. Equity/gender
4. Management

Compare this with our draft purpose statement – *To deliver relevant and high-quality free basic education to all Gambians* – and we see that quality, access and equity are all specifically mentioned. Good management is implicitly required to deliver basic education effectively and efficiently, and has been identified in the SWOT analysis of Appendix 5 as a key weakness.

We therefore have four clearly identified areas where specific improvements are required in order to deliver a programme of high-quality education. Let us now look at each one in turn to see what the key outputs might be.

2.8. Output 1 - Quality

Section 3.1 describes the challenges facing the ministry. School examinations show very poor results in core subjects. Teachers are untrained, unmotivated and as a result unprofessional in their attitude. There are no national standards and benchmarks (that

makes sense, but standards and benchmarks have to be defined to measure quality). As a result, the MTP seeks to increase competency levels in core subjects, increase contact hours, and develop a number of quality assurance and standards frameworks. And the Introduction to the MTP describes extra training and support for teachers. The whole package of activity is designed to *improve learning outcomes* (page 13).

For the purpose of the logframe, we need to formulate an output statement which tries to encapsulate the desire for quality, and this is now straightforward:

Output 1: Better quality teaching and teaching materials (resulting in improved learning outcomes)

Note that in the final logframe, the text in brackets would not appear as part of the Output statement. Improved learning outcomes will be the indicators for monitoring the effect of better quality teaching.

2.9. Output 2 - Access

Section 3.2 refers to access in terms of classroom construction and rehabilitation, to cope with the expected increase in enrolment and the reduction of drop-outs. [There is another facet of access in the MTP which is to ensure that both boys and girls have equal opportunities for basic education, but this is covered by gender.]

Section 1.2.2 also refers to access to basic education, measuring success by gross enrolment. On page 24 and 25, the MTP describes planned improvements to facilities, teaching quality and management.

Module 4 : Case studies (1) : designing logframes to improve monitoring

For the purpose of defining an appropriate output statement, we could consider a phrase which encapsulates an increase in enrolment and a decline in drop-out due to better provision of facilities and the suitability of those facilities.

Output 2: Comprehensive provision of classrooms and equipment in rural and urban areas (resulting in improved enrolment and completion of basic education)

Again, note that the text in brackets would not appear in the final logframe as part of the output statement. Improved enrolment and completion will be a measure of better access.

2.10. Output 3 – Equity/gender

Section 3.3 describes the affirmative action taken under the ESSP in order to increase girl's enrolment, retention and performance. This was successful, but had the unexpected outcome of reducing boys' enrolment. The present strategy aims for a child-friendly school initiative, but also wants to continue to make schools girl-friendly, and to ensure that the curriculum and the teachers are more sensitive to the needs of girls.

There is some overlap between these activities and those related to quality and access. Therefore it would seem sensible to concentrate here on the policies, training and materials required for ensuring gender sensitive basic education. Other aspects can be introduced by modifying the statements for outputs 1 and 2.

We might then have:

Output 1: Better quality teaching and teaching materials (resulting in improved learning outcomes for boys and girls)

Output 2: Comprehensive provision of classrooms and equipment in rural and urban areas (resulting in improved enrolment and completion of basic education for both boys and girls)

Output 3: Effective use of improved standards and policies for gender-sensitive basic education

2.11. Output 4 - Management

Underpinning all the above, but too important not to have its specific reference, is the management of basic education. The MTP gives it a separate section (4) and refers to it throughout the document. Unfortunately the document is not specific on activities, but it states that the reason for improving management "is to ensure effective and efficient delivery of education services". This will be done through service level agreements and targets for delivering objectives. There is also a reference earlier in the MTP to improved management at the school committee level.

In the real world, the people developing the logframe would have a dialogue with project personnel to examine this area more fully, and to jointly agree activities and targets, and how they might be monitored. For the purpose of this exercise, we shall have to use a less than satisfactory output statement, as follows:

Output 4: Improved management of the education system.

2.12. Draft Logical Framework for the MTP

Putting all the above together, we now have the following draft logframe for the MTP.

Super-Goal	Reduction of poverty in The Gambia
Goal	To create a human resource base in The Gambia with the education, training and skills required by VISION 2020
Purpose	To deliver relevant and high-quality free basic education to all Gambians
Output 1	Better quality teaching and teaching materials (resulting in improved learning outcomes for boys and girls)
Output 2	Comprehensive provision of classrooms and equipment in rural and urban areas (resulting in improved enrolment and completion of basic education for both boys and girls)
Output 3	Effective use of improved standards and policies for gender-sensitive basic education
Output 4	Improved management of the education system

It will be seen that the logframe by no means summarises the whole of the MTP. It should, however, be sufficient to provide a framework for developing a monitoring system, and does include all the most crucial aspects of the strategy.

3. Health Sector Reform Programme: Strategic Thrusts with a Logical Framework and a Plan of Action, 2004 – 2007, Federal Ministry of Health, Abuja

The original document is Reference D. There is relatively little text which requires annotating, so the discussion which follows will refer directly to chapters, pages and paragraph numbers for ease of reference.

3.1. Opening comments

The document needs careful study. Despite the title, it contains 9 logical frameworks, not one, for the components of the strategy, but lacks an overall unifying logframe bringing the components together. There are a number of places where the component logframes overlap slightly, and the hierarchy of the logframes provided implies that all the components are of equal "level", whereas some could be said to support others.

In the Executive Summary, amplified in the Introduction, there is a statement that the HSRP establishes a framework, including goals, targets and priorities, that should guide the action and work of the health ministry. This is true to some extent, but only for individual strategic areas where in fact the highest level (the performance objective) describes the Purpose rather than the Goal.

For the purpose of this module, it will be a useful exercise to do the following:

1. Attempt to create a unifying logical framework for the HSRP as a whole
2. Attempt to revise the 9 logical frameworks included, using the recommended format of module 1B, with Goal and Purpose, and see whether they can be fitted into a hierarchic structure of logframes.

3.2. A unifying logframe for the HSRP - discussed

The overall aims of the HSRP are summarised on page 13, paragraph 16. The HSRP will be a "process of sustained change ..... designed to improve the performance of the health system to attain better health for the Nigerian population". Paragraphs 19 and 20 then set out the Vision and Mission of the Federal Ministry of Health, which map directly onto the HSRP aims as follows:

FMOH	HSRP
Vision: to reduce morbidity and mortality, reverse prevalence of non-communicable disease, meet global targets, increase life expectancy and quality of life	Goal: To attain better health for the Nigerian population
Mission: to develop and implement policies and programmes that will strengthen the national health system to deliver effective quality and affordable health	Purpose: To improve the performance of the health system
	Outputs: fundamental changes in a number of specified areas

It is clear then that the Purpose of the HSRP is to improve the performance of the health system, and the Goal is to improve the health of Nigerians. If we require a super-goal for the HSRP, then we can look at NEEDS (see Reference E page 11), which is the over-arching development strategy for Nigeria of which the HSRP is a part.

In part 2 of the NEEDS document (Federal Government of Nigeria, 2004, part 2 Empowering People) there is the following statement:

The goal of the NEEDS health component is to improve the health status of Nigerians in order to reduce poverty..... The strategy will involve comprehensive health sector reform, aimed largely at strengthening the national health system and enhancing the delivery of effective, efficient, good quality and affordable health services.

For a super goal, therefore, we must refer to poverty reduction.

3.3. Outputs of the HSRP (1) – as stated

We turn now to the Outputs of the HSRP. We recall from Module 3 that it is usual practice to have 3 or at most 4 outputs in the logframe, to enable focused management and monitoring. This is not obligatory, but does prompt us to ask whether all the outputs mentioned in the HSRP are of the same level, or whether in fact some support others.

The HSRP contains seven strategic areas of work, as follows:

1. Improving the stewardship role of Government
2. Strengthening the national health system and its management
3. Reducing the burden of disease
4. Improving availability of health resources and their management
5. Improving access to quality health services
6. Improving consumers' awareness and community involvement
7. Promoting effective partnerships, collaboration and coordination

These seven will be sustained by two further areas for attention

8. Monitoring and evaluation, including a health management information system
9. Communication strategy

3.4. Outputs of the HSRP – examined further

Number 2 requires further examination. "Strengthening the National Health System and its Management" seems at first glance to repeat the Purpose statement "To improve the performance of the health system". What are the differences? Logframe 2 on pages 46 to 49 specifies the lines of activity for work area 2. These are: a) repositioning the Federal Ministry of Health for better performance; b) establishing a National Hospitals Agency; c) establishing new management structures for FHIs (Federal Health Institutions); d) establishing a national blood transfusion service; and e) developing a blueprint for reforming the PHC system

Module 4 : Case studies (1) : designing logframes to improve monitoring

(primary health care). This could be summarised as Planning and Co-ordination at federal level. The purpose statement refers to the performance of the entire health system, at federal and state levels, in the public and private sector.

Number 4, when the logframe on pages 56 to 58 is studied carefully, is about creating strategies and systems to support the provision of health care. This could be summarised as Resource Management. Number 5 (pages 59 to 64) is actually about Quality Control systems – not about access to health care. Number 6 (pages 65 to 67) is about Awareness and Involvement of consumers and communities, and again would not on its own have a direct impact on the Purpose.

Of the first 7, therefore, it would seem that 1, 2, 3 and 7 could be regarded as key to the achievement of the purpose. They can be summarised as Stewardship, Planning, Disease Burden and Partnerships.

3.5. A unifying logframe for the HSRP - proposed

We suggest then that the logframe for the HSRP should contain 4 outputs, and a possible version could look like this.

Super-goal (poverty)	To reduce poverty in Nigeria
Goal (better health)	To improve the health status of Nigerians
Purpose (better performance and delivery)	To improve the performance of the Nigerian health system (and deliver effective, efficient, good quality and affordable health services)
Output 1 (stewardship)	The stewardship role of Government is optimised
Output 2 (planning)	Federal planning and management is made fit for purpose
Output 3 (disease)	Reduced burden of disease
Output 4 (partnerships)	Partnerships, collaboration and coordination are made effective

3.6. A hierarchy of logframes (1) – analysed

We now turn to the detailed logframes on pages 42 to 77. Given the limitations of space on an A4 page, it is useful to summarise the activities by one or two-word phrases, in order to be able to demonstrate how logframes can be nested and made hierarchical.

For example, the Stewardship role of Government (pages 43 to 45) has what are described as four project objectives, which can be summarised as:

a) Policies   (to review, update and harmonise national health policies)
b) Legislation   (to enact a National Health Act)
c) Communication   (to deploy communication technologies for transformation, efficiency and productivity)
d) Strategies   (preparation of public health sector medium term expenditure framework strategic plans)

In a similar vein, Planning and Management (pages 46 to 49) can be summarised as:

a) Reform
b) Coordination
c) Management
d) Blood transfusion
e) PHC strategies

d), establishing a national blood transfusion system, is a supporting activity rather than an output, and might better be included as part of 3, helping to reduce of the disease burden.

Continuing, reduction of the disease burden (pages 50 to 55) has four components:

a) PHC services
b) EOC services
c) HIV+ response
d) Lifestyles

And finally, Partnerships (pages 68 and 69) can be summarised:

Module 4 : Case studies (1) : designing logframes to improve monitoring

a) Policy
b) Efficiency
c) Coordination

3.7.   A hierarchy of logframes (2) – summarised

Putting these together into a single table, using the shorthand above, we see now how the

logframes for the four main components can link with the high-level logframe.

	HSRP	Stewardship	Planning	Disease	Partnerships
Super-goal	To reduce poverty in Nigeria	To improve the health status of Nigerians	To improve the health status of Nigerians	To improve the health status of Nigerians	To improve the health status of Nigerians
Goal	To improve the health status of Nigerians	To improve the performance of the Nigerian health system	To improve the performance of the Nigerian health system	To improve the performance of the Nigerian health system	To improve the performance of the Nigerian health system
Purpose	To improve the performance of the Nigerian health system (and deliver effective, efficient, good quality and affordable health services)	To optimise the stewardship role of Government	To make Federal planning and management fit for purpose	To reduce the burden of disease	To make partnerships, collaboration and coordination more effective
Output 1	The stewardship role of Government is optimised	Policies	Reform	PHC services	Policy
Output 2	Federal planning and management is made fit for purpose	Legislation	Coordination	EOC services	Efficiency
Output 3	Reduced burden of disease	Communication	Management	HIV+ response	Coordination
Output 4	Partnerships, collaboration and coordination are made effective	Strategies	PHC strategies	Lifestyles	-

It is now a relatively straightforward task to create suitable output statements for the lower

level logframes, drawing on the text in the HSRP logframes mostly from the box labelled

"impact/outcome". This is left to the reader. These output statements will of course appear as

Activity statements on the high-level logframe.

3.8.   A hierarchy of logframes (3) – a logical tree

Regarding the logframes for strategic areas 4, 5, 6, 8 and 9, these can be done in a similar way. For example, reading the detailed text for number 5, quality control, it could be seen to be an activity which contributes to the improvement of PHC services, and in turn to reducing the disease burden. We could then have a four-level hierarchy as follows:

	HSRP	Disease	PHC services	Quality control
Super-goal	To reduce poverty in Nigeria	To improve the health status of Nigerians	To improve the performance of the Nigerian health system	To reduce the burden of disease
Goal	To improve the health status of Nigerians	To improve the performance of the Nigerian health system	To reduce the burden of disease	To improve primary health care services
Purpose	To improve the performance of the Nigerian health system	To reduce the burden of disease	To improve primary health care services	To quality control and legalise the supply of drugs
Output 1	The stewardship role of Government is optimised	Improved PHC service delivery	Needs assessment for PHC facilities	System for quality assurance designed and implemented
Output 2	Federal planning and management is made fit for purpose	Increase use of modern health facilities for essential obstetric care	Upgrading of PHC facilities	Medicinal plant species preserved and a National Pharmacopoeia established
Output 3	Reduced burden of disease	Effective health sector response to HIV/AIDS and other major diseases established	Selected schools of PHC technology refurbished	Legal drug distribution enabled and quality controlled
Output 4	Partnerships, collaboration and coordination are made effective	Healthy lifestyles promoted	Arrangement for supply of drugs implemented	-

3.9.   Summary

In the second exercise we have shown how something of the nature of a Strategy, mainly involving structures, legislation, roles and responsibilities, can still have a logical framework. In addition, for the more complex strategies, it is possible to have hierarchies of logframes

which correctly position outputs, in particular, in the logical tree leading to the highest level goals and super-goals.

The aim, of course, is to use these logframes to develop monitoring systems which correctly measure what is happening. This will be covered in modules 5, 6 and 7.

References

A   Republic of The Gambia Education Sector Medium Term Plan 2008-2011

B   Republic of The Gambia Education Sector Medium Term Plan 2008-2011 –

annotated

C   Republic of The Gambia Education Sector Strategic Plan 2006-2015

D   Health Sector Reform Programme: Strategic Thrusts with a Logical Framework and a

Plan of Action, 2004 – 2007, Federal Ministry of Health, Abuja

E   Meeting Everyone's Needs: National Economic Empowerment and Development

Strategy, Nigerian National Planning Commission, Abuja 2004 (volume 2)

Better Monitoring for Better Evaluation

Contents

1. Introduction and aims

The aim of this module is to create an awareness of the strengths and weaknesses of a number of typical indicators, including those used to monitor the millennium development goals.

Participants will learn to judge the "quality" of an indicator, and become aware of how to use both quantitative and qualitative indicators for the benefit of fact-based information and decision making. Participants will learn how to attribute appropriate indicators to the different phases of planning and implementation

In Unit 2 we analyse the OECD definition of an indicator and the various ways in which agencies and others have tried to make indicators as good as possible. In Unit 3 we discuss qualitative indicators in some depth, and their advantages and disadvantages compared with quantitative ones. Qualitative indicators – in other words non-numeric ones – are becoming increasingly important due to the fact that development projects are moving more and more upstream, and dealing with issues such as behavioural change, management effectiveness, the impact of strategy, which do not lend themselves so easily to traditional "counting".

In Unit 4 we focus on the Millennium Development Goal indicators, arguably the most high-profile set of indicators in the development world. We use the Jomtien World Declaration on Education for All as an example to show how a decision made by a world conference became a millennium development goal, and from that how international targets and indicators for educational development were agreed. This Unit contains also the current official list of millennium development goals, targets and indicators. Contrary to popular belief, the list is not static and is regularly reviewed as better information becomes available, or the results of more research are taken into account.

Module 5: Indicators: selecting and using them for better monitoring

Unit 5 is based mainly on the UNDP Handbook for Monitoring and Evaluation, which contains many useful examples of indicators for different levels of a project or programme intervention. We looked in Module 3 at logical frameworks and the importance of accurate wording for Goal, Purpose and Output: here we try to choose the most suitable indicators.

Indicators have other important characteristics apart from being qualitative or quantitative. In Unit 6 we examine four topics in greater depth: leading indicators, proxy indicators, baseline data and the setting of targets. This leads into Unit 7 where we consider data sources. We first look at the administrative outputs of the national statistical system and then describe a number of surveys in common use. We argue that national statistical systems need to do more to provide good quality monitoring data from regular sources rather than from special surveys. Finally we look briefly at other international indicators such as the Human Development Index and discuss their strengths and limitations.

2. What is an indicator?

2.1. *OECD definition*

There is no single definition of indicator, but we shall use the one provided by OECD-DAC in their *Glossary of Key terms in Evaluation*:

An indicator is "a quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect changes connected to an intervention, or to help assess the performance of a development actor".

There are a number of important things to note in this definition. First, an indicator may be quantitative – a headcount, a percentage, a rate of change – or qualitative – a ranking, an opinion.

Second, a indicator it is noted for its simplicity and reliability – we shall return to these below when discussing the SMARTness of indicators. And third, it measures achievement and reflects changes which can – through that indicator – be associated with a particular intervention or course of action; or, it can help to assess performance – in other words, it is not normally sufficient on its own for performance assessment, but needs to be taken together with a number of other factors.

A typical indicator might then be *net primary enrolment rate*. Note that concepts of time and expectations are not part of the definition. *Net primary enrolment rate increases from 60% to 70% between 2010 and 2020* is a combination of indicator, baseline and target. It is good practice to keep these three things separate, as will be discussed in Unit 6.

2.2. *Development indicators*

Indicators are signposts of change along the path to development. They describe the way to track intended results and are critical for monitoring and evaluation.

Good performance indicators are a critical part of the results framework. In particular, indicators can help to:

- Inform decision making for ongoing programme or project management
- Measure progress and achievements, as understood by the different stakeholders
- Clarify consistency between activities, outputs, outcomes and impacts
- Ensure legitimacy and accountability to all stakeholders by demonstrating progress
- Assess project and staff performance

There is a key underlying message which it is important to stress when we are looking for indicators suitable for monitoring development programmes. In the same way that the indicator light on a car tells the motorist behind what the car in front is doing (or about to do), or the indicator on a heart rate monitor makes a noise when the rate reaches a dangerous level, then a monitoring indicator must appear in a timely fashion. There is little point in knowing two years after the event that a wrong turn was taken, or that the patient was dying, when it is now too late to take corrective action.

This in turn implies that a monitoring indicator may not necessarily be the best measure available, if perfection means that information collected during 2011 is only available for use at the end of 2013. We shall see later how administrative data sources, in their attempt to cover the whole population (in its statistical sense – all the people, or all the schools, or all the deaths by cause) – may produce results just too slowly for monitoring purposes. With automation and a careful use of sampling this does not have to be the case: but in many countries the administrative systems are unable to produce results quickly enough.

Surveys are then a prime source of data, and in Unit 7 we shall discuss the main survey instruments.

2.3. *SMART or CREAMy?*

A common acronym which is used to describe the characteristics of a "good" indicator is SMART. It was originally applied to formulating objectives and targets and unfortunately it is now so often used that there are a number of different interpretations of what it actually means. Wikipedia has the following table, for example:

Letter	Major term	Other terms
S	Specific	Significant, stretching, simple
M	Measureable	Meaningful, motivational, manageable
A	Attainable	Appropriate, achievable, agreed, assignable, actionable, action-oriented, ambitious, aligned, aspirational
R	Relevant	Realistic, results-focused, results-oriented, resourced, rewarding
T	Time-bound	Time-oriented, time-framed, timed, time-based, timeboxed, timely, time-specific, timetabled, time limited, trackable, tangible

and comments that choosing certain combinations of the labels can cause duplication (attainable and realistic) or significant overlapping (appropriate and relevant). Note that according to good practice, *Time-bound* is not a characteristic of a good indicator definition, whereas *Timely* is.

Kusek and Rist in Ten steps to a results-based monitoring and evaluation system: a handbook for development practitioners (World Bank 2004) use the more useful acronym CREAM

Clear   Precise and unambiguous

Relevant   Appropriate to the subject at hand

Economic   Available at a reasonable cost

Adequate   Provides a sufficient basis to assess performance

Monitorable   Amenable to independent validation

2.4. *SPICED and more*

Roche (2002) quoted by MDF Consultancy of the Netherlands argues that when indicators are used more as specific examples of change, then different characteristics become important. He refers to SPICED indicators:

Subjective	Informants have a special position or experience that gives them unique insights which may yield a very high return on the investigators time. In this sense, what may be seen by others as 'anecdotal' becomes critical data because of the source's value.
Participatory	Indicators should be developed together with those best placed to assess them. This means involving a project's ultimate beneficiaries, but it can also mean involving local staff and other stakeholders.
Interpreted and communicable	Locally defined indicators may not mean much to other stakeholders, so they often need to be explained.
Cross-checked and compared	The validity of assessment needs to be cross-checked, by comparing different indicators and progress, and by using different informants, methods, and researchers.
Empowering	The process of setting and assessing indicators should be empowering in itself and allow groups and individuals to reflect critically on their changing situation.
Diverse and disaggregated	There should be a deliberate effort to seek out different indicators from a range of groups, especially men and women. This information needs to be recorded in such a way that these differences can be assessed over time.

Note that these characteristics are much more specific and tend to categorize the indicators according to their future use, also relating them to special ways of monitoring (eg participatory).

As we argue in Module 3, the crucial thing is to first agree on the goal, purpose and outcome statements. Clarity at this stage makes it a great deal easier to choose appropriate indicators. The principal characteristics of an indicator we require are then:

1. Specific, clear, precise and unambiguous definitions; which are
2. Measureable and reliable;

3. Relevant and appropriate to the subject;
4. Available at a reasonable cost and in time to be useful; and
5. Sufficient to provide evidence of performance

Please construct your own acronym!

2.5. *Choosing an indicator – a checklist of questions*

Indicators will be useful for monitoring any part of the chain from inputs, through activities, outputs and purpose to the goal, but should always relate directly to the result being measured. It is useful to run through the following checklist of questions for each indicator.

- How can we measure whether or not the expected results are being achieved?

- What type of information would best show change in a situation, positive or negative?

- What is it feasible to monitor with the allocated resource and capacity constraints?

- Will the indicators provide information at the right time to feed into the various monitoring programmes?

- What will the data collection system be, and who will be responsible for it?

- Can national systems be used or augmented, and are existing government indicators suitable?

2.6. *Paris Declaration and Accra Agenda*

The last point on the previous page about using national systems is particularly important given the Paris Declaration on Aid Effectiveness (2005). Paragraph 17 states that donors commit to:

17. Using a country's own institutions and systems [..] by strengthening the partner country's sustainable capacity to develop, implement and account for its policies to its citizens and

parliament. Country systems and procedures [...] include [...] results frameworks and monitoring.

This was re-emphasised by the Accra Agenda for Action (2008) where paragraph 23 reads

23. We will improve our management for results by taking the following actions:

a) Developing countries will strengthen the quality of policy design, implementation and assessment by improving information systems, including, as appropriate, disaggregating data by sex, region and socioeconomic status.

b) Developing countries and donors will work to develop cost-effective results management instruments to assess the impact of development policies and adjust them as necessary. We will better co-ordinate and link the various sources of information, including national statistical systems, budgeting, planning, monitoring and country-led evaluations of policy performance.

c) Donors will align their monitoring with country information systems. They will support, and invest in strengthening, developing countries' national statistical capacity and information systems, including those for managing aid.

The first place to look, therefore, for a suitable indicator, should be among the outputs of already established national systems. If a suitable indicator can be developed by improving those systems, then this should be given serious consideration before putting in place an entirely new and separate system.

2.7. *Choosing an indicator – other issues (1)*

We saw in Module 3 that setting Goal and Purpose statements was best done though a stakeholder consultation. This is doubly important when setting indicators, since agreement on the means of judging progress is more likely to lead to acceptance of the outcome, particularly if progress is less than satisfactory.

It is good practice to have a few good indicators rather than many poor ones. We saw in Module 3 that DFID, in their logframe guidance, suggest no more than two for Goal and Purpose, and three for each Output. This is good advice but should not be slavishly followed. There must be indicators in sufficient number to measure the breadth of the changes happening and to allow for cross-checking. These could be supplementary to those published in the logframe.

The key to good indicators is credibility—not volume of data or precision in measurement. Large volumes of data can confuse rather than bring focus and a quantitative observation is no more inherently objective than a qualitative observation. An indicator's suitability depends on how it relates to the result it intends to describe.

As far as possible, indicators should be disaggregated. Averages tend to hide disparities, and recognizing disparities is essential for programming to address the special needs of groups such as women, indigenous groups and marginalized populations. Indicators can be disaggregated by sex, age, geographic area and ethnicity, among other things.

2.8. *Choosing an indicator – other issues (2)*

Given the various tensions between cost, availability, relevance, etc, it is safe to say that indicators should direct attention to what is <u>critical</u> to the programme and whether or not it is succeeding. For example, we could be trying to measure an outcome related to a greater commitment by government partners to concerns over HIV/AIDS. A basic indicator is

Number of government ministries that have an HIV/AIDS sector strategy

but this tells us very little. A strategy could have been designed with no stakeholder involvement, no senior management engagement and no budget. Simply counting the number of ministries that have done this would not be a measure of real progress against the outcome that deals with the real commitment of the government partners. We could improve the indicator by adding

Number of government ministries that have an HIV/AIDS sector strategy developed in consultation with non-governmental stakeholders

which adds a dimension of ownership. But we can be more judgemental than this, and formulate an indicator which measures not only the existence of a strategy but also the likelihood that it will be effective.

Number of government ministries that have a strong HIV/AIDS sector strategy, where "strong" is measured by scoring (to some agreed method) against a) the strategy was developed in consultation with non-government stakeholders; b) senior officials were involved in strategy development and implementation; c) there is a budget in place to implement the strategy.

3. Indicators – quantitative or qualitative?

3.1. *Introduction*

In this Unit we examine in more depth the main issues concerning the quan/qual debate. It is often seen as an either/or situation, and both sides of the debate have strong support, usually among economists and social scientists respectively. It is wrong, however, to see this as a competition between numerical and non-numerical data. The key issue is suitability: is the indicator fit for purpose? And that means, as we have said before, what precisely is the purpose?

History is on the side of quantitative indicators, since much monitoring theory is derived from manufacturing processes where things are measured or counted – the weight of a tin, the number of items in a box. And it is true that in the early years of development assistance the focus was on activity and much of that activity was directly measureable – schools renovated, kilometres of road built, people trained. The Millennium Development Goal indicators are almost entirely quantitative (we shall study these in Unit 4).

But development practitioners now want to know about impact – the people were trained, but what did they do with their new skills? The schools were renovated, but did this improve educational outcomes? The roads were built, but did this improve rural livelihoods? And with this desire to know more, the interest has increased in qualitative measures.

3.2. *Definitions*

Quantitative data express a certain quantity, amount or range. Usually, there are measurement units associated with the data, for example the height of a person in metres. It makes sense to set boundary limits to such data, and it is also meaningful to apply arithmetic operations to the data.

Quantitative indicators are therefore classical statistical measures that measure results in terms of:

- Number (population aged 15-24 years, number of agricultural workers)
- Percentage (primary school enrolment as a percentage of children of school age)
- Rate (birth rate – births per 1,000 population)
- Ratio (sex ratio – number of males per number of females)

Qualitative data describe the attributes or properties that an object possesses. The properties are categorized into classes that may be assigned numeric values. However, there is no significance to the data values themselves; they simply represent attributes of the object concerned. Qualitative data which use clearly defined categories such as nationality, or sex, or preference, can be summarised in a quantitative way – for example, number of women of a certain nationality, or percentage of people preferring bottled water.

Qualitative indicators, therefore, are generally used to reflect people's judgements, opinions, perceptions and attitudes towards a given situation or subject. They can include changes in sensitivity, satisfaction, influence, awareness, understanding, attitudes, quality, perception, dialogue or sense of well-being. They are typically associated with terms such as *compliance with*, or *quality of*, or *extent of*, or *preference for*.

3.3. *Advantages and disadvantages of quantitative indicators*

Quantitative indicators remain popular for many good reasons. Since they are numerical measures they can be manipulated in a number of ways to produce statistical measures

such as means and medians. It is possible to analyse statistical relationships between them. It is meaningful to talk about the percentage increase in immunisation; it is less exact to talk about the change in the percentage of the population who now consider the security situation to be grave.

National statistical systems produce quantitative data as a matter of course. Most survey instruments are designed for this (and will be discussed in Unit 7). There is usually much less room for definitions and categories to be misunderstood.

However, the fact that quantitative indicators are readily available and easily understood means that there is a tendency to use them inappropriately. They are often weak indicators as they merely communicate that something has happened but not whether what has happened is an important measure of the objective, or what the quality and impact of the policies has been.

3.4. *Qualitative indicators*

Using a qualitative indicator implies that an assessment must be made of some aspect of the progress of the programme which cannot be measured using traditional statistical methods. Perhaps a project seeks to create a safer environment for women. Statistics of crimes against women and sex-disaggregated crime data will become available in due course. But perhaps focus group discussions show that, on the whole, women feel safer since the project began. This of course has many potential problems, not least in deciding whether or not the opinions are truly representative. A careful use of sampling can ensure that the groups interviewed can be said to reflect the views of all.

Another qualitative assessment which is often made involves a ranking or verbal comparison between options or outcomes, or a statement of preferences. A is preferable to B; F is better

than E; X is the best option, then Z then Y. This is a subjective view, which depends closely on the feelings of the person on the day the question was put, so the answers must be treated with caution. But they do give indications or signposts that attitudes have changed.

A third indicator involves a self-assessment of a particular variable. "Following exposure to the training programmes, how would you describe your level of skills in X?"

3.5. *Measuring Quality of Life*

Development projects are often linked to poverty reduction, and as we shall see in Unit 4, the over-arching millennium development goal is the elimination of extreme poverty and hunger. The indicators by which progress towards this will be measured are quantitative and include the proportion of the population living on less than one dollar a day, adjusted for differences between currencies and prices.

There has long been a desire to express poverty in other terms, to reflect issues such as satisfaction with life, quality of life, security, freedom, stability. Poverty is rarely an absolute concept anyway – in most countries it is based on a "minimum standard of living", as we shall see in Module 6 – and so there is plenty of room for non-monetary and indeed non-numerical standards to be incorporated.

The data collection community, however, has a long way to go before it can be said to be comprehensively measuring quality of life. One organisation which has made an effort is Gallup, and in their World Poll Database 2010, reported by the UNDP in table 9 of the 2010 Human Development Report, there are four qualitative measures of interest.

The first is *Overall life satisfaction*: 0 is least, 10 is most. The actual question is "All things considered, how satisfied are you with your life as a whole these days? Use a 0 to 10 scale where 0 is dissatisfied and 10 is satisfied."

The second is *Satisfaction with personal dimensions of well-being* and is measured by a satisfied/not satisfied response about Job satisfaction, Personal health and Standard of living.

The third is *Elements of happiness* – the percentage answering Yes to two questions: one on having a purposeful life and one on being treated with respect. The actual questions are "Do you feel your life has an important purpose?" and "Now, think about yesterday, from the morning until the end of the day. Think about where you were, what you were doing, who you were with, and how you felt. Were you treated with respect all day yesterday?"

The fourth is *Negative experience network*: 0 is most, 100 is least. This is the sum of the answers (1 for Yes and 0 for No) to 6 separate questions whether yesterday the respondent experienced worry, sadness, stress, boredom, depression and anger. Individual scores are averaged for each country or region, and rescaled to a number between 0 and 100.

3.6. *Examples of qualitative indicators from the 2010 Human Development Report*

The UNDP 2010 Human Development Report contains some instructive examples of qualitative indicators. Note the definitions of the various states of the indicator; and the source reference for those who wish to follow it up more closely.

Table 6 assesses aspects of empowerment. Of interest to us here are three examples:

Module 5: Indicators: selecting and using them for better monitoring

1. Political freedom – democracy
2. Accountability – democratic decentralisation
3. Civil liberties – press freedom

Political freedom – democracy is scored 0, 1 or 2, using the "Democracy and Dictatorship revised dataset" [https://netfiles.uiuc.edu/cheibub/www/DD_page.html]. 0 is undemocratic, 1 is democratic, but with no change in governing party, and 2 is fully democratic. The score simplifies to a great extent a large and complex dataset in the reference document, which also explains the assumptions which have been made for what is a surprisingly subjective topic.

Accountability – democratic decentralisation is similarly scored 0, 1 or 2. A score of 0 indicates there are no local elections, 1 indicates that the legislature is elected but the executive is appointed, and 2 indicates that both legislature and executive are locally elected. Details can be found in the Database of Political Institutions.

In contrast, *Civil liberties – press freedom* is an open-ended index where perfect freedom of the press scores 0 and an increasing score indicate decreasing freedom. This indicator is produced by Reporters without Borders [http://en.rsf.org/press-freedom-index-2009,1001.html] and is formed from scoring the answers to 43 questions. Many are Yes/No answers, but some require judgment on the part of the respondent. For example, question 23 asks if there was widespread self-censorship in the privately owned media and asks for a score between 0 (none) to 5 (a great deal).

3.7. *Gender and sex*

It is recommended that statistical data are disaggregated by sex – male/female – wherever possible, and in fact this is one of the requirements of the MDG indicators. However, many organisations are now interested in the *gender* attributes of the data.

Gender equality – that is people having the same chances in life, whatever their sex – is the main thrust of many development programmes, though it is more often linked to the empowerment of women. Why, for example, do fewer girls attend school than boys? Why are there fewer female parliamentarians? Traditional quantitative indicators rarely give this perspective, and qualitative methods are brought into play. It is now common to use interviews with women, for example, to examine <u>why</u> there is a sex differential in a particular statistic, or <u>how</u> women are excluded from certain activities.

DFID, in their logframe guidance, for example, insist that gender is systematically taken into account at all stages of project design and implementation. When it comes to formulating objectives, they comment that it should be clear which activities and outputs are targeted to women, which to men, and which to both. Phrases such as "the poor" should be replaced by "poor men and women". The purpose and goal statements should include promoting gender equity or women's empowerment. Indicators therefore need to be sensitive to gender.

DFID add: "Qualitative indicators are particularly important in this respect. It is not sufficient to know that women are participating in an activity: the quality of their participation and experience – whether as members of parliament, as pupils in a primary school class, or as users of public services – is all-important."

4. Millennium Development Goal indicators

4.1. *Introduction and terminology*

Among the plethora of indicators around the globe, the most well known and most often quoted are those for the MDGs. Unfortunately, because of their high profile and importance, they often eclipse other indicators which may be more relevant to specific development programmes. We shall argue that the MDGs are useful and necessary, but not always sufficient, for monitoring purposes.

We shall start by looking at the Millennium Declaration which launched the MDGs, then use Education for All to illustrate how a set of indicators were developed to measure the education goal. We then present the full current list of goals, targets and indicators and finish with a discussion of key messages.

It is necessary to be clear about terminology. The millennium development goals are verbal statements expressing an intent to improve some situation. The goals were refined into more specific targets defining the key aspects of the situation to be improved. Progress towards these targets is to be monitored by indicators.

The best document for understanding the MDGs is Indicators for Monitoring the Millennium Development Goals – Definitions, Rationale, Concepts and Sources, UN 2003.

4.2. *The Millennium Declaration*

It was the Millennium Declaration in 2000 which launched the initiative. In paragraphs 19 and 20, Governments declared that:

19. We resolve further:
• To halve, by the year 2015, the proportion of the world's people whose income is less than one dollar a day and the proportion of people who suffer from hunger and, by the same date, to halve the proportion of people who are unable to reach or to afford safe drinking water.
• To ensure that, by the same date, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling and that girls and boys will have equal access to all levels of education.
• By the same date, to have reduced maternal mortality by three quarters, and under-five child mortality by two thirds, of their current rates.
• To have, by then, halted, and begun to reverse, the spread of HIV/AIDS, the scourge of malaria and other major diseases that afflict humanity.
• To provide special assistance to children orphaned by HIV/AIDS.
• By 2020, to have achieved a significant improvement in the lives of at least 100 million slum dwellers as proposed in the "Cities Without Slums" initiative.

20. We also resolve:
• To promote gender equality and the empowerment of women as effective ways to combat poverty, hunger and disease and to stimulate development that is truly sustainable.
• To develop and implement strategies that give young people everywhere a real chance to find decent and productive work.
• To encourage the pharmaceutical industry to make essential drugs more widely available and affordable by all who need them in developing countries.
• To develop strong partnerships with the private sector and with civil society organizations in pursuit of development and poverty eradication.
• To ensure that the benefits of new technologies, especially information and communication technologies, in conformity with recommendations contained in the ECOSOC 2000 Ministerial Declaration, are available to all.

Source: http://www.un.org/millennium/summit.htm

The bullet points above form the basis for the Millennium Development Goals. The components of this declaration are of course themselves derived from a series of global summits on specific issues. Annex 4 of *Indicators for monitoring the millennium development goals* lists 21 of these, including the 1985 World Conference to Review and Appraise Achievements of the United National Decade for Women, the 1990 World Conference on Education for All, the 1992 World Conference on Environment and Development and the 1995 World Summit for Social Development.

It was the realisation that many of the declarations from these summits could not be assessed or monitored – halfway and even final reviews were demonstrating that for many countries the current situation was not known, nor even the starting point – which prompted development statisticians to push for a quantification of the goals.

4.3. *Education for All (1) – the 1990 Jomtien Declaration*

Module 5: Indicators: selecting and using them for better monitoring

We can take as an example the 1990 World Conference on Education for All. At the end of this conference in Jomtien, a declaration was made (the full text is in Annex A) of which these are the key points.

ARTICLE I - MEETING BASIC LEARNING NEEDS

Every person - child, youth and adult - shall be able to benefit from educational opportunities designed to meet their basic learning needs.

ARTICLE II - SHAPING THE VISION

To serve the basic learning needs of all requires more than a recommitment to basic education as it now exists. What is needed is an "expanded vision" that surpasses present resource levels, institutional structures, curricula, and conventional delivery systems while building on the best in current practices.

ARTICLE III - UNIVERSALISING ACCESS AND PROMOTING EQUITY

Basic education should be provided to all children, youth and adults.

ARTICLE IV - FOCUSING ON LEARNING

Whether or not expanded educational opportunities will translate into meaningful development - for an individual or for society - depends ultimately on whether people actually learn as a result of those opportunities, i.e., whether they incorporate useful knowledge, reasoning ability, skills, and values.

ARTICLE V - BROADENING THE MEANS AND SCOPE OF BASIC EDUCATION

Learning begins at birth. The main delivery system for the basic education of children outside the family is primary schooling. The basic learning needs of youth and adults are diverse and should be met through a variety of delivery systems. All available instruments

and channels of information, communications, and social action could be used to help convey essential knowledge and inform and educate people on social issues.

ARTICLE VI - ENHANCING THE ENVIRONMENT FOR LEARNING

Learning does not take place in isolation. Societies, therefore, must ensure that all learners receive the nutrition, health care, and general physical and emotional support they need in order to participate actively in and benefit from their education.

ARTICLE VII - STRENGTHENING PARTNERSHIPS

National, regional, and local educational authorities have a unique obligation to provide basic education for all, but they cannot be expected to supply every human, financial or organizational requirement for this task. New and revitalized partnerships at all levels will be necessary:

ARTICLE VIII - DEVELOPING A SUPPORTIVE POLICY CONTEXT

Supportive policies in the social, cultural, and economic sectors are required in order to realize the full provision and utilisation of basic education for individual and societal improvement.

ARTICLE IX - MOBILIZING RESOURCES

If the basic learning needs of all are to be met through a much broader scope of action than in the past, it will be essential to mobilize existing and new financial and human resources, public, private and voluntary.

ARTICLE X - STRENGTHENING INTERNATIONAL SOLIDARITY

Meeting basic learning needs constitutes a common and universal human responsibility. It requires international solidarity and equitable and fair economic relations in order to redress existing economic disparities.

Note the comprehensive nature of the Declaration, covering all aspects of education including financing and equality of access a well as quality of teaching and promoting an environment conducive to learning.

4.4. *Education for All (2) – as reflected in the Millennium Declaration*

By 2000, the Millennium Declaration had summarised the EfA declaration more succinctly and had made primary education the focus.

To ensure that, by [2015], children everywhere, boys and girls alike, will be able to complete a full course of primary schooling and that girls and boys will have equal access to all levels of education.

By 2003, the statisticians and education specialists had agreed on a slightly modified wording for two Goals, two associated Targets, and seven Indicators, which are

Goal 2: Achieve Universal Primary Education

Target 3: Ensure that, by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling

Indicator 6: Net enrolment in primary education

Indicator 7: Proportion of pupils starting grade 1 who reach grade 5

Indicator 8: Literacy rate of 15-24 year-olds

[A footnote stated that an alternative indicator – primary completion rate – was under development, and this was in fact later included as indicator 7A].

Module 5: Indicators: selecting and using them for better monitoring

Goal 3: Promote Gender Equality and Empower Women

Target 4: Eliminate gender disparity in primary and secondary education, preferably by 2005, and in all levels of education no later than 2015

Indicator 9: Ratio of girls to boys in primary, secondary and tertiary education

Indicator 10: Ratio of literate women to men, 15-24 years old

Indicator 11: Share of women in wage employment in the non-agricultural sector

Indicator 12: Proportion of seats held by women in national parliament

Note that indicators 10, 11 and 12 relate to the outcomes of a better and more equitable education for women, rather than to a simple measure of access.

4.5. *Education for All (3) – net enrolment ratio in primary education, a well-defined indicator*

In *Indicators for monitoring the Millennium Development Goals* there is detailed discussion on the indicators, why these were chosen and others were rejected. Here is the text for indicator 6. We have included it as an example of an ideal description of an indicator. Note the clear definition, reference to data sources, a discussion of the limitations of the indicator, and notes on gender.

6. NET ENROLMENT RATIO IN PRIMARY EDUCATION

DEFINITION

Net primary enrolment ratio is the ratio of the number of children of official school age (as defined by the national education system) who are enrolled in primary school to the total population of children of official school age. *Primary education* provides children with basic

reading, writing, and mathematics skills along with an elementary understanding of such subjects as history, geography, natural science, social science, art and music.

GOAL AND TARGET ADDRESSED

Goal 2. Achieve universal primary education

Target 3. Ensure that, by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling

RATIONALE

The indicator is used to monitor progress towards the goal of achieving universal primary education, identified in both the Millennium Development Goals and the Education for All initiative. It shows the proportion of children of primary school age who are enrolled in primary school. Net enrolment refers only to children of official primary school age. (Gross enrolment includes children of any age.) Net enrolment rates below 100 per cent provide a measure of the proportion of school age children who are not enrolled at the primary level. This difference does not necessarily indicate the percentage of students who are not enrolled, since some children might be enrolled at other levels of education.

METHOD OF COMPUTATION

The indicator is calculated as the number of enrolled students within the appropriate age cohort according to school records as reported to ministries of education, divided by the number of children of primary school age.

DATA COLLECTION AND SOURCE

Data on school enrolment are usually recorded by the country ministry of education or compiled from surveys and censuses. Data on the population in the official age group for the primary level are available from national statistical offices, based on population censuses

and vital statistics registration. Nationally reported values will be the same as internationally reported values only if the same methods and population estimates are used.

For international comparisons and estimates of regional and global aggregates, the UNESCO Institute for Statistics regularly produces data series on school enrolment based on data reported by education ministries or national statistical offices and United Nations population estimates. For countries for which administrative data are not available, household survey data may be used to assess school attendance rather than enrolment. Among international surveys, the Multiple Indicator Cluster Survey and Demographic and Health Surveys (and sometimes Living Standards Measurement Surveys and the Core Welfare Indicators Questionnaires in Africa) provide school attendance data.

PERIODICITY OF MEASUREMENT

Enrolment data are recorded regularly by ministries of education and are available on a yearly basis. Data derived from surveys and censuses, when administrative records on enrolment by age and sex are not available, are less frequent. Net enrolment rates produced by UNESCO are available on an annual basis for two thirds of countries, but usually one year after the reference year. The United Nations Population Division estimates population by individual years of age biannually, although estimates may be based on population censuses conducted every 10 years in most countries. Household survey data, such as those from the Multiple Indicator Cluster Survey and Demographic and Health Surveys, are available for many developing countries at regular intervals of three to five years.

GENDER ISSUES

In situations of limited resources, families make difficult choices about sending their children to school. They may perceive the value of education differently for boys and girls. Girls are more likely than boys to suffer from limited access to education, especially in rural areas. Nevertheless, where basic education is widely accepted and overall enrolment is high, girls tend to equal or outnumber boys at primary and secondary levels.

DISAGGREGATION ISSUES

Rural and urban differences are particularly important in the analysis of enrolment data owing to significant differences in school facilities, available resources, demand on children's time for work and dropout patterns. It is also important to consider disaggregation by geographical areas and social or ethnic groups. However, showing and analysing data on specific ethnic groups may be a sensitive issue in the country. Gender differences may also be more pronounced in some social and ethnic groups.

COMMENTS AND LIMITATIONS

School enrolments may be over-reported for various reasons. Survey data may not reflect actual rates of attendance or dropout during the school year. Administrators may report exaggerated enrolments, especially if there is a financial incentive to do so. Children who repeat years may mistakenly be included in the net figures. Children's ages may be inaccurately estimated or misstated. Census data may be out of date or unreliable. There may also be insufficient data on school enrolment by sex, but existing measurement problems make it difficult to assess the situation correctly. The indicator attempts to capture the education system's coverage and efficiency, but it does not solve the problem completely. Some children fall outside the official school age because of late or early entry rather than because of grade repetition. Enrolment data compiled by UNESCO are adjusted to be consistent with the International Standard Classification of Education, 1997 (ICSCED) and are therefore comparable across countries. National data derived from administrative records are not necessarily based on the same classification over time and may not be comparable with data for other countries, unless exactly the same classification is used. Similarly, the concepts and terms in household surveys and censuses do not necessarily remain constant over time.

REFERENCES AND INTERNATIONAL DATA COMPARISONS

ASIAN DEVELOPMENT BANK. Gender Issues in Basic and Primary Education. In *Gender Checklist: Education*. Manila. Available from http://www.adb.org/documents/Manuals/ Gender_checklists/Education.

UNITED NATIONS (1998). *Principles and Recommendations for Population and Housing Censuses, Revision 1*. Series M, No. 67, Rev. 1, para. 2.156 Sales No. E.98.XVII.1. Available from http://unstats.un.org/unsd/pubs (A, E, F, S).

UNITED NATIONS (2003). *Millennium Indicators Database*. Statistics Division Internet site http://millenniumindicators.un.org.

UNITED NATIONS CHILDREN'S FUND (2003). *Monitoring Methods*. New York. Internet site http:// unicef.org/reseval/methodr. html.

UNITED NATIONS CHILDREN'S FUND (annual). *The State of the World's Children*. New York.

UNITED NATIONS DEVELOPMENT PROGRAMME (2003 and annual). *Human Development Report*. New York: Oxford University Press. Available from http://hdr.undp.org .

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION (1978). *Revised* Recommendations Concerning the International Standardization of Educational Statistics. Paris. See also UNESCO Statistical Yearbook, 1998, chap.2.

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION (1997). *International* Standard Classification of Education, 1997(ISCED). Montreal. Available from http://www.uis.unesco.org. Path: Core Theme Education, Technical Guides.

WORLD BANK (2003 and annual). *World Development Indicators*. Print and CD-ROM. Washington, D.C. Available in part from http://www.worldbank.org/data .

UNESCO data since 1998 follow the 1997 version of the International Standard Classification of Education, 1997 ISCED, which enables international comparability between countries. The time series data before 1998

are not consistent with data for 1998 and after.

AGENCIES

Ministries of Education

UNESCO Institute for Statistics

Source: Indicators for monitoring the Millennium Development Goals, 2003

4.6. *Discussion*

There are a number of important issues related to the MDG indicators and indicators in general, which the Education for All example illustrates.

Do not try to invent the perfect indicator when a satisfactory one already exists, particularly if it is internationally accepted and understood, is already collected nationally as a matter of course, or is part of an international survey instrument.

But do not overlook the limitations of an indicator, and be prepared to replace it with a better one in time. This is one of the most difficult aspects of the MDGs to appreciate. Most development agencies and governments now have MDG indicators built into their performance plans, and wish for consistency so they can monitor progress from year to year. This has raised the profile of the MDG indicators, and in some way reduced the profile of the actual Goals. It seems sometimes to be more important to meet the indicator target than to achieve the Goal. It is of course the Goals which should be built into development plans: these are what most countries have signed up to achieve.

Module 5: Indicators: selecting and using them for better monitoring

Do remember that one indicator can be used to measure progress in a number of goals. In the case above, indicator 10 is a direct measure of gender equality and an outcome measure of education for all.

Definitions which are implicit in the indicator may not be universally accepted. For example, primary education does not cover the same age group in every country. Thanks, however, to the efforts of national statistical offices in liaison with the UNESCO Institute for Statistics (UIS) in Montreal, differences between country definitions are now much more clearly understood and it is possible, using UNESCO data, to make some meaningful comparisons. It is true that in the MDG document cited, this is clearly explained – but few people refer to or recall the detailed notes.

The MDGs, when first developed, were intended to broadly measure a country's progress, and to allow for comparisons between countries. In fact some of them are only calculated at national level, and some, indeed, are only produced by international agencies. [The $1 a day headcount and the maternal mortality rate are two examples]. In the last decade we have seen people trying to apply them to districts or even towns, which in most cases is asking too much of the indicator (and of the data, and of the statisticians).

Finally, users of the MDGs should note the general comments in the introduction. The key one is:

Five main criteria guided the selection of indicators. Indicators should:

1. Provide relevant and robust measures of progress towards the targets of the Millennium Development Goals
2. Be clear and straightforward to interpret and provide a basis for international comparison
3. Be broadly consistent with other global lists and avoid imposing an unnecessary burden on country teams, Governments and other partners
4. Be based to the greatest extent possible on international standards, recommendations and best practices
5. Be constructed from well-established data sources, be quantifiable and be consistent to enable measurement over time

This is very good advice for all indicators. The text goes on to say:

Just as the indicator list is dynamic and will necessarily evolve in response to changing national situations, so will the metadata change over time as concepts, definitions and methodologies change.

Source: Indicators for Monitoring the Millennium Development Goals, page 1, UN 2003

The intended dynamism of the indicator list has been largely forgotten.

4.7. *The Millennium Development Goals, Targets and Indicators*

Effective 15 January 2008

Official list of MDG indicators	
All indicators should be disaggregated by sex and urban/rural as far as possible	
Goals and Targets (from the Millennium Declaration)	Indicators for monitoring progress
Goal 1: Eradicate extreme poverty and hunger	
Target 1A: Halve, between 1990 and 2015, the proportion of people whose income is less than one dollar a day	1.1 Proportion of population below $1 (PPP) per day (1) 1.2 Poverty gap ratio 1.3 Share of poorest quintile in national consumption
Target 1B: Achieve full and productive employment and decent work for all including women and young people	1.4 Growth rate of GDP per person employed 1.5 Employment-to-population ratio 1.6 Proportion of employed people living below $1 (PPP) per day 1.7 Proportion of own-account and contributing family workers in total employment
Target 1C: Halve, between 1990 and 2015, the proportion of people who suffer from hunger	1.8 Prevalence of underweight children under five years of age 1.9 Proportion of population below minimum level of dietary energy consumption
Goal 2: Achieve universal primary education	
Target 2A: Ensure that, by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling	2.1 Net enrolment ratio in primary education 2.2 Proportion of pupils starting grade 1 who reach last grade of primary 2.3 Literacy rate of 15-24 year olds, women and men
Goal 3: Promote gender equality and empower women	
Target 3A: Eliminate gender disparity in primary and secondary education, preferably by 2005, and in all levels of education by 2015	3.1 Ratios of girls to boys in primary, secondary and tertiary education 3.2 Share of women in wage employment in the non-agricultural sector 3.3 Proportion of seats held by women in national parliament
Goal 4: Reduce child mortality	
Target 4A: Reduce by two-thirds, between 1990 and 2015, the under-five mortality rate	4.1 Under-five mortality rate 4.2 Infant mortality rate 4.3 Proportion of 1 year old children immunised against measles
Goal 5: Improve maternal health	
Target 5A: Reduce by three quarters, between 1990 and 2015, the maternal mortality ratio	5.1 Maternal mortality ratio 5.2 Proportion of births attended by skilled health personnel

Target 5B: Achieve, by 2015, universal access to reproductive health	5.3 Contraceptive prevalence rate
	5.4 Adolescent birth rate
	5.5 Antenatal care coverage (at least one visit and at least four visits)
	5.6 Unmet need for family planning

(1) For monitoring country poverty trends, indicators based on national poverty lines should

be used, where available

Official list of MDG indicators	
All indicators should be disaggregated by sex and urban/rural as far as possible	
Goals and Targets (from the Millennium Declaration)	Indicators for monitoring progress
Goal 6: Combat HIV/AIDS, malaria and other diseases	
Target 6A: Have halted by 2015 and begun to reverse the spread of HIV/AIDS	6.1 HIV prevalence among population aged 15-24 years
	6.2 Condom use at last high-risk sex
	6.3 Proportion of population aged 15-24 years with comprehensive knowledge of HIV/AIDS
	6.4 Ratio of school attendance of orphans to school attendance of non-orphans aged 10-14 years
Target 6B: Achieve, by 2010, universal access to treatment for HIV/AIDS for all those who need it	6.5 Proportion of population with advanced HIV infection with access to antiretroviral drugs
Target 6C: have halted by 2015 and begun to reverse the incidence of malaria and other major diseases	6.6 Incidence and death rates associated with malaria
	6.7 Proportion of children under 5 sleeping under insecticide-treated bednets
	6.8 Proportion of children under 5 with fever who are treated with appropriate anti-malarial drugs
	6.9 Incidence, prevalence and death rates associated with tuberculosis
	6.10 Proportion of tuberculosis cases detected and cured under directly observed treatment short courses
Goal 7: Ensure environmental sustainability	
Target 7A: Integrate the principles of sustainable development into country policies and programmes and reverse the loss of environmental resources	7.1 Proportion of land area covered by forest
	7.2 CO2 emissions, total, per capita and per $1 GDP (PPP)
	7.3 Consumption of ozone-depleting substances
	7.4 Proportion of fish stocks within safe biological limits
Target 7B: Reduce biodiversity loss, achieving, by 2010, a significant reduction in the rate of loss	7.5 Proportion of total water resources used
	7.6 Proportion of terrestrial and marine areas protected
	7.7 Proportion of species threatened with extinction
Target 7C: Halve, by 2015, the proportion of people without sustainable access to safe drinking water and basic sanitation	7.8 Proportion of population using an improved drinking water source
	7.9 Proportion of population using an improved sanitation facility
Target 7D: By 2020, to have achieved a significant improvement in the lives of at least 100 million slum dwellers	7.10 Proportion of urban population living in slums (2)

(2) The actual proportion of people living in slums is measured by a proxy, represented by

the urban population living in households with at least one of the four characteristics: (a) lack

of access to improved water supply; (b) lack of access to improved sanitation; (c)

overcrowding (3 or more persons per room); and (d) dwellings made of non-durable material.

Official list of MDG indicators	
All indicators should be disaggregated by sex and urban/rural as far as possible	
Goals and Targets (from the Millennium Declaration)	Indicators for monitoring progress
Goal 8: develop a global partnership for development	
Target 8.A: Develop further an open, rule-based, predictable, non-discriminatory trading and financial system	Some of the indicators listed below are monitored separately for the least developed countries (LDCs), Africa, landlocked developing countries and small island developing States.
Includes a commitment to good governance, development and poverty reduction – both nationally and internationally	Official development assistance (ODA)
	8.1 Net ODA, total and to the least developed countries, as percentage of OECD/DAC donors' gross national income
Target 8.B: Address the special needs of the least developed countries	8.2 Proportion of total bilateral, sector-allocable ODA of OECD/DAC donors to basic social services (basic education, primary health care, nutrition, safe water and sanitation)
Includes: tariff and quota free access for the least developed countries' exports; enhanced programme of debt relief for heavily indebted poor countries (HIPC) and cancellation of official bilateral debt; and more generous ODA for countries committed to poverty reduction.	8.3 Proportion of bilateral official development assistance of OECD/DAC donors that is untied
	8.4 ODA received in landlocked developing countries as a proportion of their gross national incomes
	8.5 ODA received in small island developing States as a proportion of their gross national incomes
	Market access
	8.6 Proportion of total developed country imports (by value and excluding arms) from developing countries and least developed countries, admitted free of duty
Target 8.C: Address the special needs of landlocked developing countries and small island developing States (through the Programme of Action for the Sustainable Development of Small Island Developing States and the outcome of the twenty-second special session of the General Assembly)	8.7 Average tariffs imposed by developed countries on agricultural products and textiles and clothing from developing countries
	8.8 Agricultural support estimate for OECD countries as a percentage of their gross domestic product
	8.9 Proportion of ODA provided to help build trade capacity
	Debt sustainability
	8.10 Total number of countries that have reached their HIPC decision points and number that have reached their HIPC completion points (cumulative)
Target 8.D: Deal comprehensively with the debt problems of developing countries through national and international measures in order to make debt sustainable in the long term	8.11 Debt relief committed under HIPC and MDRI Initiatives
	8.12 Debt service as a percentage of exports of goods and services
Target 8E: In cooperation with pharmaceutical companies, provide access to affordable essential drugs in developing countries	8.13 Proportion of population with access to affordable essential drugs on a sustainable basis
Target 8F: In cooperation with the private sector, make available the benefits of new technologies, especially information and communications	8.14 Telephone lines per 100 population
	8.15 Cellular subscribers per 100 population
	8.16 Internet users per 100 population

The Millennium Development Goals and targets come from the Millennium Declaration, signed by 198 countries, including 147 Heads of State and Government, in September 2000 (http://www.un.org/millennium/declaration/ares552e.htm) and from further agreement by member states at the 2005 World Summit (Resolution adopted by the General Assembly – A/RES/60/1, http://www.un.org/Docs/journal/asp/ws.asp?m=A/RES/60/1). The goals and targets are interrelated and should be seen as a whole. They represent a partnership between the developed countries and the developing countries "to create an environment – at the national and global levels alike – which is conducive to development and the elimination of poverty".

5. Examples of Goal, Purpose and Output indicators

5.1. *Scope*

In this Unit we shall concentrate on indicators at the Goal, Purpose and Output level, in the same way that this was the focus of Module 3 when we discussed goal, purpose and output statements in the context of logical frameworks. The reason is that monitoring inputs and activities is relatively straightforward and mainly involves simple measures such as quantities of materials, expenditure or time.

We recall from Module 3 that the Goal of a programme is associated with Impact and the Purpose is associated with an Outcome, in the usual terminology. But in the varied literature on monitoring we shall find equal references to goal-level indicators and impact indicators, and to purpose-level indicators and outcome indicators.

The Asian Development Bank, in their *Guidelines on preparing a design and monitoring framework* has this useful summary.

Impact-level Indicator. Performance indicators at the impact level specify the expected medium to long-term impact at program, subsector, or sector level to which the present project, several other projects, and initiatives described in the CSP will contribute. Hence, the impact level indicator includes targets beyond the scope of the project. It is essential that the responsibility for monitoring and measuring the impact indicators is clearly defined during project processing.

Outcome-level Indicator. The outcome or end-of-project indicator defines the project's immediate effect on or the behavioural change of beneficiaries, and improvements to systems or institutions. They are the performance targets that the project takes full accountability to deliver and they are the basis by which the project will be judged a success or a failure.

Output-level Indicator. Output indicators are the easiest to define. They specify the key tangible goods and services the project will deliver. They define the project management's terms of accountability that have to be achieved by the end of project implementation. It is the project management's responsibility to monitor the performance indicators at output level.

We might argue with the statement that output indicators are the easiest to define, if our output, say, is a functioning electoral system, but it is generally accepted that Outputs themselves are the more straightforward of the three to accurately express in words. In fact, the output itself may be measurable and serve as its own indication of whether or not it has been produced.

We shall now look at each in turn, with examples and comments. All the examples are taken from the UNDP Handbook on Monitoring and Evaluation, 2009. The comments are ours.

5.2. *Goal-level/Impact indicators*

The goal, it will be remembered, is shared with government and all the other actors in the sector concerned. For sufficiently high level development programmes, the goal is often related to achieving one or more of the MDGs, and so very often it is appropriate to use MDG indicators for measuring impact – perhaps specifying a geographic or social group, depending on the programme.

It will be necessary to consider the frequency with which indicators such as these are likely to be available, and whether they will appear often enough to tell us if the programme is going off-track. It could be argued that when considering the wider goal, it is not so necessary to have (for example) annual impact indicators, as long as the purpose level indicators appear in time for us to take any action required.

5.3. *Example 1*

Goal: Increased public participation in national and local elections, particularly by women, indigenous populations and other traditionally marginalised groups.

Possible indicators:

- Proportion of eligible voters who vote in the national (or local) elections
- Percentage of eligible women voters who vote in the elections
- Percentage of eligible indigenous people who vote in the elections

Comments: These are good quantitative indicators which need a solid statistical base. Census information will give the eligible voting population – certainly by sex, but not always defining the indigenous population. Voting turnout is usually done as a headcount at polling stations: we need to ask whether this will identify members of indigenous groups. Gender specialists will also want qualitative information on the reasons why women voters did not vote, in time to modify any affirmative actions of the programme.

5.4. *Example 2*

Goal: Improved educational performance of students in region A

Possible indicators:

- Percentage of students completing primary schooling

- Pass rates in standardised tests

Comments: The goal is very wide and embraces, as it is currently stated, all levels of education. Pass rates in standardised tests (in region A) is therefore a good indicator of performance. The percentage completing primary schooling is open to discussion. If we assume that primary schooling lasts 6 years then the project is some years old before children are completing who have benefitted from having most of their schooling under the improved system.

In addition, therefore, it might be useful to monitor intake to primary, secondary and tertiary education – which should increase if the quality of education is seen to be improving.

All these indicators should be disaggregated by sex.

5.5. *Example 3*

Goal: Reduction in poverty and hunger

Possible indicators:

- Poverty rate
- Gini coefficient
- Percentage of population living in extreme poverty
- Level of infant malnutrition

Comments: The goal is close to MDG1, and some might wish it to be more specific, saying what kind of reduction is expected. Many countries are now saying their development aim is to *Achieve the MDG targets for poverty and hunger*, in other words reduce poverty and hunger to half of their 1990 level by 2015. Despite the inclusion of time and quantity, this is a valid goal. It means, however, that these two targets will be pursued to the exclusion of what

may be better strategies (a cynic might say lifting people to just above the poverty line in time to be counted out, and then abandoning them).

The indicators are similar to the MDG1 indicators, and need no particular comment, apart from to say that they are rarely available more often than once a year, and usually only once every two or three years. It might also be useful to monitor subsidiary indicators, such as the numbers of children requiring support from a school midday meal programme, or reported deaths from starvation, or migration in search of food and work.

5.6. *Purpose-level/Outcome indicators – Example 4*

Purpose: Electoral administrative policies and systems reformed to ensure freer and fairer elections and to facilitate participation by marginalised groups

Possible indicators:

- Percentage of citizens surveyed who believe that the electoral management process is free and fair
- Percentage of women and minorities surveyed that are aware of their rights under the new electoral administration laws
- Annual percentage increase in the number of women registered to vote
- Annual percentage increase in the number of indigenous people registered to vote
- Ratio of voter registration centres to population in rural areas

Comments: The first two indicators are good measures of outcome, but will require special surveys to collect the information. Funding for these surveys needs to be built into the project. The data can be collected by carefully run opinion polls; indicators of fairness could be found in press reports, if the press is considered to be unbiased. Some elections are run under the aegis of international observers, whose opinions are relevant.

Voter registration is a key indicator but the central records system needs to be checked to ensure that it is possible to identify sex and membership of indigenous groups. Otherwise, a

special survey will need to be mounted to estimate this indicator and whether the increase in registration of women, for example, is better or worse than that for men.

5.7. *Example 5*

Purpose: To improve the e-governance capacity of key central government ministries and agencies by 2015

Possible indicators:

- The extent to which key central government bodies have strong on-line facilities for citizen engagement. This is a composite indicator which covers
    o Key central government ministries have websites established
    o Websites contain functional contact information
    o Websites contain functional access to major government policy documents and publications
    o Websites facilitate access to persons with disabilities (or is available in a second language)
    o Websites provide links to other major government departments
    o Websites facilitate online payments for important government services (taxes, motor vehicle registration, etc)
- Percentage of property tax revenue collected through online payment system

Comments: This is a very interesting example of the use of combined quantitative and qualitative data which can be collected with little effort. The establishment of a website is a yes or no question. Functional access could mean access was operational on the day it was tested; but should mean is generally accessible more often than not. Provision of links to other government departments is one issue – but that these links are operational and kept up to date also needs to be recorded since web addresses change. The facilitation of on-line payments is a clear definition.

What is missing, perhaps, is a measure of connectivity of the population – in other words their ability to take advantage of e-government.

5.8. *Example 6*

Purpose: Reduction in the level of violence against women by 2013

Possible indicators:

- Number of reported cases of domestic abuse against women
- Percentage of women who feel that violence against women has reduced within the last five years
- Proportion of men who believe that wife beating is justified for at least one reason

Comments: The first indicator should come from regular crime statistics, if they exist, but only if the violence is reported. It is a partial but official quantitative measure. The other two indicators are derived from surveys and in fact these questions are now part of the Demographic and Health Survey questionnaire (see Unit 7 for more detail). The data analysis will be good but infrequent. It may be necessary to set up special surveys to collect the data if a DHS is not planned.

5.9. *Output indicators – Example 7*

Output: National electoral management agency has systems, procedures and competencies to administer free and fair elections.

Possible indicators:

- Percentage of electoral centres using multiple forms of voter identification measures
- Number of centres that are headed by trained professional staff
- Percentage of electoral management office staff and volunteers trained in techniques to reduce voter fraud
- Percentage of electoral management office staff who believe that their agency is more professional and better run than one year ago

Comments: The indicators suggested go most of the way to monitoring the output, but it would also be of interest to know whether the electoral management agency has an open and transparent structure which will allow the systems and procedures to be operated

without interference and for problems to be reported and solved. Trained staff are necessary but they also need to be able to put their training into practice without fear of intimidation.

5.10. *Example 8*

Output:　　　National human development report produced and disseminated

Possible indicators:

- Number of copies of report distributed
- Percentage of parliamentarians who receive a copy
- Extent to which national human development report findings and recommendations were used to inform high-level policy discussions

Comments:　　This is quite a weak output statement and in other circumstances would be an activity which contributed towards better-informed policy discussions. The number of reports distributed is a fact, but not very informative. Is it better or worse, for example, to produce 1000 copies this time compared to 750 the time before? We would really want to be sure that the report was sent to all the key players; that they received it; and that they read it and shared it with colleagues. A short questionnaire accompanying the covering letter could perhaps inspire some comments from recipients.

The third indicator is not a measure of "production and dissemination" but of impact, and is much more important.

5.11. *Example 9*

Output:　　　Civil society and community organisations in region have resources and skills to contribute to monitoring of local poverty reduction strategies.

Module 5: Indicators: selecting and using them for better monitoring

Possible indicators:

- Number of NGO staff completing training courses in poverty analysis by end of year
- Percentage of trained NGO staff who feel that they are more effective at doing their jobs one year later
- Percentage of districts with Monitoring Committees
- Percentage of districts with Citizen Community Boards

Comments: This is a complex output. The organisations need resources and the people need skills. They need to be able to make a positive contribution to monitoring local poverty reduction strategies. The set of indicators proposed nearly covers the output. The first two could be combined: the percentage of staff who have completed training who can work more effectively as a result. The last two could also be combined, since it is merely a count. It is also necessary to judge how effective the Monitoring Committees and Community Boards are – perhaps by the outcomes of meetings – and to ensure that they are not prevented from working effectively by lack of resources.

6. Other aspects of indicators

6.1. *Introduction*

In the earlier Units we have looked at indicators suitable for monitoring impact, purpose and outputs, based on some kind of logical framework which summarises the project or programme. We have seen that indicators need certain characteristics – we have suggested

1. Specific, clear, precise and unambiguous definitions; which are
2. Measureable and reliable;
3. Relevant and appropriate to the subject;
4. Available at a reasonable cost and in time to be useful; and
5. Sufficient to provide evidence of performance

We shall now take a closer look at two issues related to bullet points 4 and 5: what do we do when the ideal indicator will come too late to be really useful; and what do we do when the ideal indicator is not available at all? In these conditions we need leading indicators and proxy indicators.

Following this, we shall consider baselines and targets, both intermediate and final.

6.2. *Leading indicators*

The term *leading indicator* is well known to national accountants and economic statisticians, who are required to provide early estimates of the direction of economic growth to markets and government, perhaps a year before the full analysis is published. This can be done by a partial analysis of partial data; or by finding an indicator which is easy to calculate and which, through experience, is a good predictor of others.

In development programmes, the situation is slightly different but the science is the same. For example, let us say we have a programme to improve educational attainment which can be measured quite well by examination results. Unfortunately, by the time we get the results from students who have benefited from the programme (as opposed to those who were already in the system) it is too late to make any changes. We therefore need additional indicators which will be available quickly and which have the potential to predict exam outcomes. Reduced drop-out rates, or increased enrolment, could be used.

Another example could be a programme to reduce infant mortality. Here, although we can expect the impact of the programme to be more immediate, it still takes some time for the data to be collected and analysed. We might set up a random sample of clinics or hospitals and analyse this partial data, or we might target clinics which have high mortality and set up systems to analyse their data more quickly.

6.3. *1 year old immunisation as a leading indicator for under-5 mortality*

In the MDG list, indicator 4.3 is the proportion of one-year old children immunised against measles, and is used to measure progress towards reducing the under-five mortality rate. The notes say:

The indicator provides a measure of the coverage and the quality of the child health-care system in the country. Immunization is an essential component for reducing under-five mortality. Governments in developing countries usually finance immunization against measles and diphtheria, pertussis (whooping cough) and tetanus (DPT) as part of the basic health package. Among these vaccine-preventable diseases of childhood, measles is the leading cause of child mortality. Health and other programmes targeted at those specific causes are one practical means of reducing child mortality.

The first dose of measles vaccine is supposed to be administered to all children at the age of nine months or shortly after. By 2000, most countries were providing a "second opportunity" for measles vaccination, either through a two-dose routine schedule or through a combined routine schedule and supplementary campaigns. Measles immunization coverage is expressed as the percentage of children who have received at least one dose. Vaccination coverage for measles needs to be above 90 per cent to stop transmission of the virus—not only because measles is so contagious, but also because up to 15 per cent of children vaccinated at nine months fail to develop immunity. Some countries in the Latin America and Caribbean region, for example, administer the measles vaccine at 12–15 months of age. This has to be taken into account in calculations of coverage based on household surveys.

In many developing countries, lack of precise information on the size of the cohort of children under one year of age makes immunization coverage difficult to estimate.

6.4. *Proxy indicators*

It is sometimes difficult to directly measure the preferred indicator, and an indirect, or proxy, indicator is required instead. These should only be used when the data for direct indicators are not available, perhaps because the cost of data collection cannot be justified, or if it is not feasible to collect direct information as frequently as desired. It may also be impossible to collect for security reasons – perhaps in certain slum or ghetto areas. But caution is needed in the use of a proxy indicator, as there must be a presumption that the proxy is giving at least an approximate view of the desired outcome.

There are many proxy indicators in use nationally and internationally, and a selection of them are described below. Some have been developed as a result of sound academic research, showing that changes in x are strongly correlated with changes in y. Some are more intuitive,

such as using bicycles or tin roofs to measure changes in rural income, or purchase of fertiliser to indicate greater agricultural production.

Similarly, in the absence of reliable data on corruption in countries, many development agencies use the information from surveys that capture the perception of corruption by many national and international actors as a proxy indicator.

In the Human Development Index, UNDP and other UN organizations use 'life expectancy' as a proxy indicator for health care and living conditions. The assumption is that if people live longer, then it is reasonable to assume that health care and living conditions have improved. Real gross domestic product/capita (purchasing power parity) is also used in the same indicator as a proxy indicator for disposable income.

6.5. *Child malnutrition as a proxy for hunger*

There are a number of proxy indicators in the MDG list and two examples are given here. Note the scientific reasoning behind the choice of proxy, and the clear demonstration that a change in the proxy can safely be used to infer a likely change in the main indicator.

MDG Indicator 1.8, prevalence of underweight children under five years of age, is used to measure progress towards halving the proportion of people who suffer from hunger. The notes say:

Child malnutrition, as reflected in body weight, is selected as an indicator for several reasons. Child malnutrition is linked to poverty, low levels of education and poor access to health services. Malnourishment in children, even moderate, increases their risk of death, inhibits their cognitive development, and affects health status later in life. Sufficient and good

quality nutrition is the cornerstone for development, health and survival of current and succeeding generations. Healthy

nutrition is particularly important for women during pregnancy and lactation so that their children set off on sound developmental paths, both physically and mentally. Only when optimal child growth is ensured for the majority of their people will Governments be successful in their efforts to accelerate economic development in a sustained way.

The under-five underweight prevalence is an internationally recognized public health indicator for monitoring nutritional status and health in populations. Child malnutrition is also monitored more closely than adult malnutrition.

6.6. *Seats held by women as a proxy for gender inequality*

MDG Indicator 3.3, proportion of seats held by women in national parliament, is used to measure progress towards eliminating gender disparity in all levels of education. The notes say:

Women's representation in parliaments is one aspect of women's opportunities in political and public life, and it is therefore linked to women's empowerment. Parliaments vary considerably in their independence and authority, though they generally engage in law-making, oversight of Government and representation of the electorate. In terms of measuring women's real political decision-making, this indicator may not be sufficient, because women still face many obstacles in fully and efficiently carrying out their parliamentary mandate. Thus, being a member of parliament, especially in developing countries and emerging democracies, does not guarantee that a woman has the resources, respect or constituency to exercise significant influence.

6.7. *Baseline data*

All agencies agree that baseline data are a necessary pre-requisite for project or programme approval, though this is not always enforced. It is common sense, however, that you must know the starting point in order to measure and monitor progress. The various global declarations referred to in Unit 4, and initiatives such as the decade for the eradication of poverty and the decade for water and sanitation, suffered greatly from a lack of data about the starting position. Promises to halve poverty between year x and year y are meaningless without a baseline.

Development projects and programmes are usually designed to respond to a problem: the background analysis can often be a fruitful source of baseline data. This will of course include qualitative data, as we saw earlier.

6.8. *Targets*

The list of indicators for monitoring a project or programme is not complete without adding to it a time-bound target plus, if the target is some years away, some interim targets. In the MDG example of Unit 4, we have:

Indicator 1.1 – proportion of the population under $1 PPP a day
Target 1A - Halve, between 1990 and 2015, the proportion of people whose income is less than one dollar a day

This requires a little extra work before it can be used. We need to know the situation in 1990 in order to set the target for 2015; we can then derive the 2015 target as 50% of the 1990 figure; and then we can set interim targets. For the $1 a day figure, it has been usual to take

a straight line approach in order to do this, though geometric or logarithmic approaches to interim targets can just as easily be justified.

In fact, setting interim targets should not be done without consultation. Say, for example, it is known that there are a significant proportion of the population who are just below the poverty line, and another cluster that are far below. One could imagine that the first group could be brought out of poverty relatively quickly; then there may be a period when nothing appears to be happening; then the second group are brought out of poverty after some effort. Interim targets which have been set by drawing a straight line from a to b might then suggest better performance at the start than is really the case.

It is equally necessary to set interim and final targets for qualitative indicators, though this requires more imagination.

7. Sources of data

7.1. *Introduction*

In module 3, we saw how most agencies suggest it is best to have two or three indicators for each level of the logical framework. We have seen earlier in this module that the MDG targets usually have three or more indicators each. This is to try to reflect different views of the target while still retaining a relative simplicity. It is sensible then to first review the indicators which are already being collected as a matter of course, by governments and agencies, to see to what extent they are useful for monitoring the strategy or programme in question.

As we have argued earlier, for high level strategic monitoring, it would be curious if the appropriate MDG indicator were not included; but it would be equally curious for an MDG indicator to be used to monitor an activity, say, which has no chance of having a direct or indirect impact on it. It must be possible to attribute some or all of any movement in the indicator to the project or programme.

This section, then, will review the main sources of data for monitoring strategies and programmes.

7.2. *Data from the National Statistical System*

The first place to look for data is among the outputs of the national statistical system (NSS), in line with the Paris Declaration and the Accra Agenda. Even if there are reservations about quality and timeliness, the scope and comprehensiveness of the data produced must make

the NSS a logical starting point. An early involvement of the statisticians in the formulation of the strategy and logframe is naturally very helpful, and will increase the chance of finding suitable indicators, and the data with which to measure them, from national sources. There is nothing more depressing (or futile) for the NSS than to be approached two years into the programme with the question "how are we progressing with X?"

Administrative data are famous for their thorough coverage of the nation, but are also notorious for the length of time they take to be collected, analysed and published. Some statistical offices produce "leading indicators" – often of the economy – which are well understood by users to be less accurate than the detail which will eventually be published, but are sufficient for the purpose. There would seem to be no insurmountable barrier to producing leading indicators of other sectors, probably by a careful use of sampling.

7.3. *Administrative sources*

The most commonly used sources of data for education, and often for health, are administrative sources – data derived from the process of administration of education or health. Unemployment data are also frequently derived from administrative registrations of employment offices. The data are made available by ministries and sometimes by national statistical offices. Data on births and deaths are also frequently drawn from administrative sources, usually vital statistics registration systems. Administrative sources can potentially provide data for very small areas. Their disadvantages include bias, application of national standards and definitions, and non-response. For vital statistics, such as births and deaths, and for many other indicators, the data often do not exist or are incomplete.

We have cited the UNESCO Institute of Statistics earlier: they provide a good example of how some lateral thinking, combined with appropriate technology, has noticeably speeded up

Module 5: Indicators: selecting and using them for better monitoring

the production of internationally comparable education statistics from standard administrative sources.

It [The UNESCO Institute for Statistics (UIS)] ... mainly focused on improving, on the one hand, quality, comparability and timeliness of primary and secondary education data gathered from countries and, on the other, data processing and transmission methods.

Recent DGF funding has played a large part in enabling substantial modernization of methods for handling the worldwide database that UIS is responsible for. Its Data Capture Initiative (DCI) uses a common technology for all surveys, produces them in different languages, and uses Adobe, which requires no special installation. An Error Reporting System (ERS) and other comparable tools have brought economies. UIS is in the process of moving from fixed-date collection campaigns to a system of Rolling Collection and Release, shortening the time between data supply and its reflection in the database. Joint use of SDMX is being developed with Eurostat and OECD.

UIS has also continued to make progress in improving country coverage of the basic statistics on primary and secondary education, their international comparability, quality and timeliness. This has been much helped by periodic meetings on a sub-regional basis with the Education Ministry staff responsible for countries' own questionnaires and surveys. In a few countries EU aid has enabled UIS to play a much larger role in redeveloping the whole reporting system.

One of the few clear improvements revealed by the extensive review of international databases that we have carried out concerns primary enrolment data for Sub-Saharan African countries: there has been a visible shift from estimated data to country data.

Source: Marrakech Action Plan for Statistics: Report on Phase 2 of an Independent

Evaluation, Christopher Willoughby and Phil Crook, World Bank 2008

7.4. *National surveys*

The National Statistical System can also provide data from surveys. Most offices conduct

household surveys on a regular basis – often of household income and expenditure, for the

purpose of price indices and for social indicators – and these should be good sources of

information. They unfortunately suffer in many countries from under-funding and slow and

inflexible analysis.

HOUSEHOLD BUDGET SURVEYS

Household budget surveys (HBS) are intended for various purposes, including measurement

of poverty and of household consumption of goods and services for weighting consumer

prices. Their value as sources of other data derives from the inclusion of a variety of

questions among basic or general variables. This offers the possibility of cross-classifying

them against many other variables, including income and urban or rural location as well as all

the common classifiers such as age and gender. The surveys are complex and expensive,

so they are not conducted very frequently in developing countries. They are undertaken often

enough (perhaps every five years or so) in many countries to provide fairly up-to-date and

fairly frequent data, however.

LABOUR FORCE SURVEYS

Labour force surveys (LFS) have become widespread in industrialized countries, but are

rarer in developing countries. They are intended to provide information on employment and

unemployment, but they also frequently seek information on education and training and may

include other variables. They occasionally include questions on income from employment. For reasons of efficiency, they generally cover the non-institutional population. In developing countries, they are often undertaken only in urban areas.

Source: Indicators for the Millennium Development Goals

7.5. *International surveys*

A number of surveys are successfully being implemented world-wide, and have been specifically designed or modified to provide many of the MDG indicators. These include the following, and are examined one by one.

- MICS – the Multiple Indicator Cluster Survey – developed by UNICEF and undertaken by many national statistical offices
- DHS – the Demographic and Health Survey – sponsored by USAID and undertaken by ICFMacro
- LSMS – the Living Standards Measurement Survey – launched by the World Bank and sponsored mainly by the WB and USAID, with other donors
- CWIQ – the Core Welfare Indicators Questionnaire – developed by the World Bank

7.6. *Multiple Indicator Cluster Survey*

The United Nations Children's Fund developed MICS in 1994 in order to provide data on some of the goals set by the 1990 World Summit for Children. This was with the realisation that standard statistical sources would not be sufficient – note that the world was half way through the decade before much of the data collection started [see the remarks above on baseline data].

MICS1 was a pilot for urgently needed mid-decade indicators: MICS2 was revised and greatly expanded. The MICS4 round is now being undertaken and is currently confirmed for

43 countries. The topics of the survey focus on children, with modules on health, development, education, child protection and HIV/AIDS.

Until now, MICS have been carried out every 5 years, but it is intended to make them 3-yearly from round 4.

More information can be found at http://www.childinfo.org.

7.7. *Demographic and Health Survey*

Demographic and Health Surveys (DHS) are sponsored by the United States Agency for International Development and undertaken by ICFMacro Inc. They were first conducted in 1984 as successors to the International Statistical Institute World Fertility Surveys. The Demographic Health Surveys are now in their fourth series. The abbreviated name was changed in 1997 to DHS+. They have been undertaken in over 60 countries. Some countries have had only one DHS, but others have had several. Most surveys are addressed to about 5,000 households.

The strength and utility of DHS is a) the focus on MDG indicators and b) the comparability between countries. Analysis is done in the United States to a common methodology.

The main areas on which DHS collects information are fertility, family planning, maternal and child health, child survival, HIV/AIDS, malaria, nutrition, women's empowerment, health inequities, domestic violence and adult health.

More information can be found at www.measuredhs.com.

7.8. *Living Standards Measurement Study*

The first Living Standards Measurement Study (LSMS) surveys were launched by the World Bank in Côte d'Ivoire and Peru in 1985. Since then there have been about 40 surveys in 25 countries. They have been sponsored by various donors, including the World Bank and the United States Agency for International Development, and by some individual countries. Although the first few LSMS surveys followed a similar format, they have varied considerably since then. There are standard LSMS modules, but they are often omitted. The organization of the fieldwork also varies. Most are one-off sample surveys, but a four wave panel was also undertaken in the Kagera region of Tanzania. More information can be found at
http://www.worldbank.org/lsms.

7.9. *Core Welfare Indicators Questionnaire*

The Core Welfare Indicators Questionnaire (CWIQ) survey is relatively new. Developed by the World Bank, it was piloted in Kenya in 1996 and in Ghana in 1997. The questionnaires are relatively short (about eight sides), but other modules may be added. The surveys are intended to be annual and to have samples of 5,000 to 15,000 households. The questionnaire is designed to complement other surveys as part of a national monitoring package. It is intended to contribute to statistical capacity-building in developing countries. The results are intended to be available within a few weeks of data collection....More information can be found at http://www4.worldbank.org/afr/stats/cwiq.cfm

7.10. *Other specialised surveys*

In addition, there are a number of very specialist surveys which are being carried out with the assistance and sponsorship of interested agencies. We shall not go into details here, but the list includes:

- HIV/AIDS, various epidemiological surveys (Joint United Nations Programme on HIV/AIDS)

- Tuberculosis/DOTS (notification programme)

- Roll Back Malaria (World Health Organization)

- Pilot surveys in selected countries to test/improve methodologies of data collection on labour force (International Labour Organization)

- Child labour survey (International Labour Organization)

- Informal sector surveys (International Labour Organization)

- Pilot surveys in selected countries to test/improve methodologies of data collection on nutrition (Food and Agriculture Organization of the United Nations)

- Pilot small-scale studies on education/literacy (United Nations Educational, Scientific and Cultural Organization)

- Access to personal computers and the Internet (International Telecommunication Union)

- Secure tenure and slum improvement (United Nations Human Settlements Programme)

Source: Indicators for monitoring the Millennium Development Goals, 2003

7.11. *Census*

Finally it is necessary to stress the importance of censuses of population and housing. Most of the MDG indicators (and other indicators which may be developed) assume that the total population of the area in which the development programme is operating, is known, whether it is the nation or a group of villages. Not only that, but the data should be available disaggregated by sex and by certain key age groups, such as under 5, 15-24, working age.

CENSUSES OF POPULATION AND HOUSING

A population census is the primary source of information about the number of people in a country and the characteristics of the population. Several features distinguish a census from survey-based sources of data. It can achieve complete coverage of the population. It offers possibilities for relating individual characteristics of the population with those of households. It provides details about subnational population groups. Owing to its high cost, it has the disadvantage of being able to provide data only once every 10 years, or sometimes less, and the questionnaires have to be relatively short. The census is the unique basic source of benchmark demographic data, such as number of people by age and gender. Demographic data are used as denominators for ratios of all kinds, on school enrolment for example, and for many other common country assessment and Millennium Development Goals indicators. However, population estimates have to be updated between censuses and national methods and standards can differ. Many international agencies use United Nations estimates of population as denominators for ratios in order to be consistent between countries. Nevertheless, the United Nations population estimates, which are revised every two years, are often different from the national estimates, mainly (but not always) as a result of international standardization. Censuses are also sometimes used as sampling frames for sample surveys.

Source: Indicators for monitoring the Millennium Development Goals, 2003

7.12. *Discussion*

There are a number of interesting points to make relating to the above.

1.   The international community consider the MDGs to be sufficiently important to invest substantial resources into developing survey instruments. Adequate funding is available to run them in developing countries just often enough to be useful – say every 3 years on

average. This means that indicators are available to show a medium term direction of change, but not really often enough for monitoring.

2.   These surveys are intended to be comparable, but definitions do change from survey to survey and care needs to be taken in reading the small print. But they have an undoubted advantage in the speed at which they can be analysed and results published.

3.   International surveys of the sort described are primarily aimed at providing good information as quickly as possible. Capacity building is a by-product, rather than a purpose. National survey programmes can be delayed or hampered by the demands of international ones, though it is true that surveys such as DHS and MICS are becoming more and more a part of national programmes.

4.   It is often overlooked, but in order to calculate the $1 a day at PPP prices indicator, a good set of national accounts are required, plus participation in the ICP – the International Comparison of Prices Programme. So despite the apparent lack of focus on economic data in the MDG list, it is inherent in the dollar a day indicator.

5.   The $1 a day international poverty measure can distract from national poverty lines (which it is intended were used for national purposes)

6.   Notwithstanding point 4, the MDG indicators are mainly social and environmental and in less well resourced offices their calculation tend to divert resources from areas such as trade, transport, industrial production, agriculture

7.   MDG indicators are selective and can encourage a too limited focus, for example on primary enrolment to the detriment of measures of education quality

8.   The fact that a survey can provide for example two measures of an indicator 3 years apart, which is sufficient to keep donors happy, can take away the impetus to improve the administrative data which could provide the data annually

7.13. *Other international data sources*

In addition to the MDG indicators, a number of international institutions have developed and promoted sets of indicators for more specialist purposes. Both the World Development Report of the World Bank and the Human Development Report of UNDP are rich sources of ideas.

For financial and monetary indicators, the IMF publishes a large number. FAO do the same for agriculture and fisheries. The UNESCO Institute of Statistics and the ILO create indicators in their respective fields. Many of these can be found in the UNDATA website http://data.un.org which conveniently brings together most of the international data sets.

It is open to discussion how much of what is published there can really be described as an indicator, rather than just a regular data series. For the purpose of this course they are unlikely to be useful for monitoring anything but the goals of the highest level strategies.

There are advantages and disadvantages of using indicators such as these. The advantages include:

- They reduce the cost of developing a purpose-built monitoring system
- They allow for comparison between similar projects and programmes
- They are quality controlled by the database owner
- They are usually standardised to meet international definitions

But there are a number of disadvantages:

- They do not necessarily address national goals
- They may not be readily accepted or owned by stakeholders

- Indicators which have been standardised to international definitions may then not be comparable with other national data
- They are unlikely to be available at sub-national level

The majority of the indicators to be found in the documents mentioned above are based on well-established statistical methodology with a firm theoretical foundation. The advantage of this is clear: definitions are well understood, sources are unbiased and non-contradictory, questionnaires and other data collection instruments are readily available, and there are long time series for all indicators.

7.14. *The Human Development Index*

Perhaps the most well known index is the UNDP Human Development Index, which was created in 1990 for the first Human Development Report to respond to criticism that the World Bank's *World Development Report* tended to compare levels of development with too much weight given to economic indicators.

The HDI is a composite index with three components, education, longevity and standard of living. It has been refined and modified over the last 20 years, and the most recent HDR 2010 describes the current calculation. The description which follows is based on the Technical Notes to HDR2010.

The Human Development Index (HDI) is a summary measure of human development. It measures the average achievements in a country in three basic dimensions of human development: a long and healthy life, access to knowledge and a decent standard of living. The HDI is the geometric mean of normalised indices measuring achievements in each dimension.

A long and healthy life is assessed by the indicator life expectancy at birth. This is normalised (converted to a figure between 0 and 1) by comparing it with a minimum life

expectancy of 20 years (chosen by UNDP) and a maximum of 83.2 (observed in Japan in 2010). [The equation is (actual life expectancy-20)/(83.2-20)]

Access to knowledge is assessed by two indicators, mean years of schooling and expected years of schooling. Both are given a minimum value of 0 years, with maxima of 13.2 and 20.6 respectively (based on observed maxima during the period 1980-2010). The two indicators are normalised and then combined into one figure using the geometric mean.

Standard of living is assessed by annual gross national income per capita at purchasing power parity, normalised between an observed minimum of $163 and observed maximum of $108,211.

This creates three figures each between 0 and 1, which are then combined using the geometric mean into a single figure, the HDI.

Bibliography

Key documents

[1] Handbook on Planning, Monitoring and Evaluating for Development Results, UNDP 2009

[2] Guidelines on Preparing a Design and Monitoring Framework, Asian Development Bank, 2005

[3] Indicators for Monitoring the Millennium Development Goals: Definitions, Rationale, Concepts and Sources, UN 2003

[4] Guidance on using the revised Logical Framework, DFID 2009

Other documents

[5] Glossary of Key Terms in Evaluation and Results Based Management, OECD-DAC 2010

[6] Ten Steps to a Results-Based Monitoring and Evaluation System, World Bank 2004

[7] World Development Report 2010, World Bank 2010

[8] Human Development Report 2010, UNDP 2010

[9] United Nations Millennium Declaration, UN 2000

[10]   The Paris Declaration on Aid Effectiveness and the Accra Agenda for Action, OECD 2005/2008

Links

[11]   Multiple Indicator Cluster Survey - www.childinfo.org.

[12]   Demographic and Health Survey - www.measuredhs.com.

[13]   Living Standards Measurement Study - www.worldbank.org/lsms.

[14]   Core Welfare Indicators Questionnaire - www4.worldbank.org/afr/stats/cwiq.cfm

[15]   World Declaration on Education for All (Jomtien Declaration) - www.unesco.org/education/efa/ed_for_all/background/world_conference_jomtien.shtml

Appendix A: Jomtien Declaration

World Declaration on Education For All

Meeting Basic Learning Needs

Preamble

More than 40 years ago, the nations of the world, speaking through the Universal Declaration of Human Rights, asserted that "everyone has a right to education". Despite notable efforts by countries around the globe to ensure the right to education for all, the following realities persist:

• More than 100 million children, including at least 60 million girls, have no access to primary schooling;

• More than 960 million adults, two-thirds of whom are women, are illiterate, and functional illiteracy is a significant problem in all countries, industrialized and developing;

• More than one-third of the world's adults have no access to the printed knowledge, new skills and technologies that could improve the quality of their lives and help them shape, and adapt to, social and cultural change; and

• More than 100 million children and countless adults fail to complete basic education programmes; millions more satisfy the attendance requirements but do not acquire essential knowledge and skills;

At the same time, the world faces daunting problems: notably mounting debt burdens, the threat of economic stagnation and decline, rapid population growth, widening economic disparities among and within nations, war, occupation, civil strife, violent crime, the preventable deaths of millions of children and widespread environmental degradation. These problems constrain efforts to meet basic learning needs, while the lack of basic education

among a significant proportion of the population prevents societies from addressing such problems with strength and purpose.

These problems have led to major setbacks in basic education in the 1980s in many of the least developed countries. In some other countries, economic growth has been available to finance education expansion, but even so, many millions remain in poverty and unschooled or illiterate. In certain industrialized countries too, cutbacks in government expenditure over the 1980s have led to the deterioration of education.

Yet the world is also at the threshold of a new century, with all its promise and possibilities. Today, there is genuine progress toward peaceful detente and greater cooperation among nations. Today, the essential rights and capacities of women are being realized. Today, there are many useful scientific and cultural developments. Today, the sheer quantity of information available in the world - much of it relevant to survival and basic well-being - is exponentially greater than that available only a few years ago, and the rate of its growth is accelerating. This includes information about obtaining more life-enhancing knowledge - or learning how to learn. A synergistic effect occurs when important information is coupled with another modern advance - our new capacity to communicate.

These new forces, when combined with the cumulative experience of reform, innovation, research and the remarkable educational progress of many countries, make the goal of basic education for all - for the first time in history - an attainable goal.

Therefore, we participants in the World Conference on Education for All, assembled in Jomtien, Thailand, from 5 to 9 March, 1990:

Recalling that education is a fundamental right for all people, women and men, of all ages, throughout our world;

Understanding that education can help ensure a safer, healthier, more prosperous and environmentally sound world, while simultaneously contributing to social, economic, and cultural progress, tolerance, and international cooperation;

Knowing that education is an indispensable key to, though not a sufficient condition for,

personal and social improvement;

Recognizing that traditional knowledge and indigenous cultural heritage have a value and validity in their own right and a capacity to both define and promote development;

Acknowledging that, overall, the current provision of education is seriously deficient and that it must be made more relevant and qualitatively improved, and made universally available;

Recognizing that sound basic education is fundamental to the strengthening of higher levels of education and of scientific and technological literacy and capacity and thus to self-reliant development; and

Recognizing the necessity to give to present and coming generations an expanded vision of, and a renewed commitment to, basic education to address the scale and complexity of the challenge; proclaim the following

World Declaration on Education for All:

Meeting Basic Learning Needs

EDUCATION FOR ALL: THE PURPOSE

ARTICLE I - MEETING BASIC LEARNING NEEDS

1. Every person - child, youth and adult - shall be able to benefit from educational opportunities designed to meet their basic learning needs. These needs comprise both essential learning tools (such as literacy, oral expression, numeracy, and problem solving) and the basic learning content (such as knowledge, skills, values, and attitudes) required by human beings to be able to survive, to develop their full capacities, to live and work in dignity, to participate fully in development, to improve the quality of their lives, to make informed decisions, and to continue learning. The scope of basic learning needs and how they should be met varies with individual countries and cultures, and inevitably, changes with the passage of time.

2. The satisfaction of these needs empowers individuals in any society and confers upon

them a responsibility to respect and build upon their collective cultural, linguistic and spiritual heritage, to promote the education of others, to further the cause of social justice, to achieve environmental protection, to be tolerant towards social, political and religious systems which differ from their own, ensuring that commonly accepted humanistic values and human rights are upheld, and to work for international peace and solidarity in an interdependent world.

3. Another and no less fundamental aim of educational development is the transmission and enrichment of common cultural and moral values. It is in these values that the individual and society find their identity and worth.

4. Basic education is more than an end in itself. It is the foundation for lifelong learning and human development on which countries may build, systematically, further levels and types of education and training.

EDUCATION FOR ALL:

AN EXPANDED VISION AND A RENEWED COMMITMENT

ARTICLE II - SHAPING THE VISION

To serve the basic learning needs of all requires more than a recommitment to basic education as it now exists. What is needed is an "expanded vision" that surpasses present resource levels, institutional structures, curricula, and conventional delivery systems while building on the best in current practices. New possibilities exist today which result from the convergence of the increase in information and the unprecedented capacity to communicate. We must seize them with creativity and a determination for increased effectiveness.

As elaborated in Articles III-VII, the expanded vision encompasses:

- Universalizing access and promoting equity;

- Focussing on learning;

- Broadening the means and scope of basic education;

- Enhancing the environment for learning;

- Strengthening partnerships.

The realization of an enormous potential for human progress and empowerment is contingent upon whether people can be enabled to acquire the education and the start needed to tap into the ever-expanding pool of relevant knowledge and the new means for sharing this knowledge.

ARTICLE III - UNIVERSALIZING ACCESS AND PROMOTING EQUITY

1. Basic education should be provided to all children, youth and adults. To this end, basic education services of quality should be expanded and consistent measures must be taken to reduce disparities.

2. For basic education to be equitable, all children, youth and adults must be given the opportunity to achieve and maintain an acceptable level of learning.

3. The most urgent priority is to ensure access to, and improve the quality of, education for girls and women, and to remove every obstacle that hampers their active participation. All gender stereotyping in education should be eliminated.

4. An active commitment must be made to removing educational disparities. Underserved groups: the poor; street and working children; rural and remote populations; nomads and migrant workers; indigenous peoples; ethnic, racial, and linguistic minorities; refugees; those displaced by war; and people under occupation, should not suffer any discrimination in access to learning opportunities.

5. The learning needs of the disabled demand special attention. Steps need to be taken to provide equal access to education to every category of disabled persons as an integral part of the education system.

ARTICLE IV - FOCUSSING ON LEARNING

Whether or not expanded educational opportunities will translate into meaningful development - for an individual or for society - depends ultimately on whether people actually learn as a result of those opportunities, i.e., whether they incorporate useful knowledge,

reasoning ability, skills, and values. The focus of basic education must, therefore, be on actual learning acquisition and outcome, rather than exclusively upon enrolment, continued participation in organized programmes and completion of certification requirements. Active and participatory approaches are particularly valuable in assuring learning acquisition and allowing learners to reach their fullest potential. It is, therefore, necessary to define acceptable levels of learning acquisition for educational programmes and to improve and apply systems of assessing learning achievement.

ARTICLE V - BROADENING THE MEANS AND SCOPE OF BASIC EDUCATION

The diversity, complexity, and changing nature of basic learning needs of children, youth and adults necessitates broadening and constantly redefining the scope of basic education to include the following components:

- **Learning begins at birth.** This calls for early childhood care and initial education . These can be provided through arrangements involving families, communities, or institutional programmes, as appropriate.

- **The main delivery system for the basic education of children outside the family is primary schooling.** Primary education must be universal, ensure that the basic learning needs of all children are satisfied, and take into account the culture, needs, and opportunities of the community. Supplementary alternative programmes can help meet the basic learning needs of children with limited or no access to formal schooling, provided that they share the same standards of learning applied to schools, and are adequately supported.

- **The basic learning needs of youth and adults are diverse and should be met through a variety of delivery systems.** Literacy programmes are indispensable because literacy is a necessary skill in itself and the foundation of other life skills. Literacy in the mother-tongue strengthens cultural identity and heritage. Other needs can be served by: skills training, apprenticeships, and formal and non-formal education programmes in health, nutrition, population, agricultural techniques, the environment, science, technology, family life,

including fertility awareness, and other societal issues.

• All available instruments and channels of information, communications, and social action could be used to help convey essential knowledge and inform and educate people on social issues. In addition to the traditional means, libraries, television, radio and other media can be mobilized to realize their potential towards meeting basic education needs of all.
These components should constitute an integrated system - complementary, mutually reinforcing, and of comparable standards, and they should contribute to creating and developing possibilities for lifelong learning.

ARTICLE VI - ENHANCING THE ENVIRONMENT FOR LEARNING

Learning does not take place in isolation. Societies, therefore, must ensure that all learners receive the nutrition, health care, and general physical and emotional support they need in order to participate actively in and benefit from their education. Knowledge and skills that will enhance the learning environment of children should be integrated into community learning programmes for adults. The education of children and their parents or other caretakers is mutually supportive and this interaction should be used to create, for all, a learning environment of vibrancy and warmth.

ARTICLE VII - STRENGTHENING PARTNERSHIPS

National, regional, and local educational authorities have a unique obligation to provide basic education for all, but they cannot be expected to supply every human, financial or organizational requirement for this task. New and revitalized partnerships at all levels will be necessary: partnerships among all sub-sectors and forms of education, recognizing the special role of teachers and that of administrators and other educational personnel; partnerships between education and other government departments, including planning, finance, labour, communications, and other social sectors; partnerships between government and non-governmental organizations, the private sector, local communities, religious groups,

and families. The recognition of the vital role of both families and teachers is particularly important. In this context, the terms and conditions of service of teachers and their status, which constitute a determining factor in the implementation of education for all, must be urgently improved in all countries in line with the joint ILO/ UNESCO Recommendation Concerning the Status of Teachers (1966). Genuine partnerships contribute to the planning, implementing, managing and evaluating of basic education programmes. When we speak of "an expanded vision and a renewed commitment", partnerships are at the heart of it.

EDUCATION FOR ALL: THE REQUIREMENTS

ARTICLE VIII - DEVELOPING A SUPPORTIVE POLICY CONTEXT

1. Supportive policies in the social, cultural, and economic sectors are required in order to realize the full provision and utitlization of basic education for individual and societal improvement. The provision of basic education for all depends on political commitment and political will backed by appropriate fiscal measures and reinforced by educational policy reforms and institutional strengthening. Suitable economic, trade, labour, employment and health policies will enhance learners' incentives and contributions to societal development.
2. Societies should also insure a strong intellectual and scientific environment for basic education. This implies improving higher education and developing scientific research. Close contact with contemporary technological and scientific knowledge should be possible at every level of education.

ARTICLE IX - MOBILIZING RESOURCES

1. If the basic learning needs of all are to be met through a much broader scope of action than in the past, it will be essential to mobilize existing and new financial and human resources, public, private and voluntary. All of society has a contribution to make, recognizing that time, energy and funding directed to basic education are perhaps the most profound investment in people and in the future of a country which can be made.

2. Enlarged public-sector support means drawing on the resources of all the government agencies responsible for human development, through increased absolute and proportional allocations to basic education services with the clear recognition of competing claims on national resources of which education is an important one, but not the only one. Serious attention to improving the efficiency of existing educational resources and programmes will not only produce more, it can also be expected to attract new resources. The urgent task of meeting basic learning needs may require a reallocation between sectors, as, for example, a transfer from military to educational expenditure. Above all, special protection for basic education will be required in countries undergoing structural adjustment and facing severe external debt burdens. Today, more than ever, education must be seen as a fundamental dimension of any social, cultural, and economic design.

ARTICLE X - STRENGTHENING INTERNATIONAL SOLIDARITY

1. Meeting basic learning needs constitutes a common and universal human responsibility. It requires international solidarity and equitable and fair economic relations in order to redress existing economic disparities. All nations have valuable knowledge and experiences to share for designing effective educational policies and programmes.

2. Substantial and long-term increases in resources for basic education will be needed. The world community, including intergovernmental agencies and institutions, has an urgent responsibility to alleviate the constraints that prevent some countries from achieving the goal of education for all. It will mean the adoption of measures that augment the national budgets of the poorest countries or serve to relieve heavy debt burdens. Creditors and debtors must seek innovative and equitable formulae to resolve these burdens, since the capacity of many developing countries to respond effectively to education and other basic needs will be greatly helped by finding solutions to the debt problem.

3. Basic learning needs of adults and children must be addressed wherever they exist. Least developed and low-income countries have special needs which require priority in

international support for basic education in the 1990s.

4. All nations must also work together to resolve conflicts and strife, to end military occupations, and to settle displaced populations, or to facilitate their return to their countries of origin, and ensure that their basic learning needs are met. Only a stable and peaceful environment can create the conditions in which every human being, child and adult alike, may benefit from the goals of this Declaration.

We, the participants in the World Conference on Education for All, reaffirm the right of all people to education. This is the foundation of our determination, singly and together, to ensure education for all. We commit ourselves to act cooperatively through our own spheres of responsibility, taking all necessary steps to achieve the goals of education for all. Together we call on governments, concerned organizations and individuals to join in this urgent undertaking. The basic learning needs of all can and must be met. There can be no more meaningful way to begin the International Literacy Year, to move forward the goals of the United Nations Decade of Disabled Persons (1983-92), the World Decade for Cultural Development (1988-97), the Fourth United Nations Development Decade (1991-2000), of the Convention on the Elimination of Discrimination against Women and the Forward Looking Strategies for the Advancement of Women, and of the Convention on the Rights of the Child. There has never been a more propitious time to commit ourselves to providing basic learning opportunities for all the people of the world. We adopt, therefore, this **World Declaration on Education for All: Meeting Basic Learning Needs** and agree on the **Framework for Action to Meet Basic Learning Needs**, to achieve the goals set forth in this **Declaration**.

Source:

http://www.unesco.org/education/efa/ed_for_all/background/world_conference_jomtien.shtml

Better Monitoring for Better Evaluation

Content

1. Introduction and Preparation

1.1.   Aim and Learning goals

The module should answer the question: What is poverty? Maybe it will raise more questions because there is no easy and simple answer to this. The earliest definitions of poverty centred on the inability to obtain adequate food and other basic necessities. Today, the main focus continues to be on material deprivations but it will be argued that this notion of economic welfare remains too narrow to reflect individual well-being. If the notion of spurring the active efforts over the past several decades to expand the concept of poverty has reached the participants then the aim has been attained.

Participants should understand the different concepts of poverty, the methods of poverty monitoring and understand and conceive the conception and implication of a poverty monitoring system and its impact on political decisions.

1.2.   Initial Words / Abstract

Poverty is multifaceted, manifested by conditions that include malnutrition, inadequate shelter, unsanitary living conditions, unsatisfactory and insufficient supplies of clean water, poor solid waste disposal, low educational achievement and the absence of quality schooling, chronic ill health, and widespread common crime. Through the signing of the Millennium Development Declaration in 2000, 191 UN member states unanimously committed to reducing poverty. However, because it is not easy to define or measure, monitoring poverty in its broad manifestations is a complex task conceptually and empirically. To get informed about the central issues of poverty and answers to the pertinent questions How much poverty is there?

Module 6: Poverty: measuring and monitoring it for development interventions

Who are the poor?

What are the characteristics of their living conditions?

How does poverty evolve over time?

please read the "HANDBOOK ON POVERTY STATISTICS: CONCEPTS, METHODS AND

POLICY USE" as of Dec 2005 by the UNITED NATIONS STATISTICS DIVISION. [1]

..or study the most important contents of this document, the UN Handbook.

Many of the aspects of this document will be dealt with in this course; some of the contents

are actually accepted as course material because there is not really a more encompassing

document on the issue of Poverty Statistics.

When it comes to the purpose of this course: At the outcome the course should answer some

questions and enable the participant to understand the concepts of poverty, the methods of

poverty monitoring and understand and conceive the conception and implication of a poverty

monitoring system and its impact on political decisions.

Some questions:

What is poverty? The earliest definitions of poverty centred on the inability to obtain

adequate food and other basic necessities. Today, the main focus continues to be on

material deprivations, i.e., the failure to command private resources. Development experts,

including [33], though, have argued that this notion of economic welfare remains too narrow

to reflect individual well-being, spurring active efforts over the past several decades to

expand the concept of poverty.[1]

---

[1] [1] Page 27

Module 6: Poverty: measuring and monitoring it for development interventions

Some basic findings:

Not surprisingly, a single, all-encompassing measure of poverty remains beyond reach. Recognizing the trade-offs between different methods of measuring poverty, researchers are now seeking compromises by integrating **qualitative** and **quantitative** indicators into their analyses.

Some basic deductions:

Despite all this, social scientists still find it useful to focus largely on poverty as a lack of money measured either as **low income** or as **inadequate expenditures**. One reason for focusing on money is practical: inadequate income is clear, measurable, and of immediate concern for individuals. Another reason is that low incomes tend to correlate strongly with other concerns that are important but harder to measure.

The case of Mozambique and in particular the province of Sofala will play a prominent part in this course for several reasons:

One of the authors have spent several years in the Statistical service on Mozambique and there in particular as a consultant for provincial services, among them the province of Sofala.

The case of poverty in Mozambique and Sofala is well documented and secondary material is readily available

More important the primary data are available for the most recent Household Surveys, Censuses and several other surveys.

Abundant analytical material on poverty in Mozambique is available for comparison and reference

Primary data for other poverty related surveys are available

It is not meant that the case of Mozambique should be an example for other countries. On the contrary there are several shortcomings but they are important because they are frequent and widespread. Mozambicans and their friends can only learn from them.

Eradication of poverty is only possible if information about poverty is available and how to do it is a task all stakeholders and governments with responsibility for the well-being of the country's people have a tremendous task to shoulder.

It will be a constant trait of this document, the user's guide or handout, to refer to practical examples. The use of more sophisticated (and costly) software will be avoided. The only commercial software used in the examples is Microsoft's EXCEL. The software usually is readily available for most users ; if not **Open Office** (http://download.openoffice.org/other.html) will be the choice of software. A shareware EXCEL Add-in statistiXL (http://www.statistixl.com/) is used for the more advanced statistical exercises. This cannot substitute completely specialized software like SPSS but serves well for the purpose of the course.

Tables, Boxes and Graphs are marked and indexed separately and usually illustrate in abstract form some special subject.

1.3.   *Some basic concepts*

1.3.1. *The poverty line*

A poverty line typically specifies the income (or level of spending) required purchasing a bundle of essential goods (typically food, clothing, shelter, water, electricity, schooling, and reliable healthcare). Identifying the poor as those with income (or expenditures) **below** a given line brings clarity and focus to policy making and analysis.

Having a poverty line allows to count the poor, target resources, and monitor progress against a clear benchmark. Communicating the extent of poverty becomes easier, and explaining the notion of deprivation simpler.

Module 6: Poverty: measuring and monitoring it for development interventions

It is also important to bear in mind differences between concepts and reality. The fact is that a poverty line (below which one is poor and above which one is not) has little empirical correspondence in the daily lives of the poor. Researchers analyzing data on households see no clear breaks or discontinuities in the relationship of income and health or nutrition, and certainly no systematic breaks in living standards that correspond to poverty lines as the term is used. Yet, poverty measures based on poverty lines serve an important descriptive purpose and should be seen in that light.

Several examples are taken from a rather recent HH Survey being conducted in Mozambique 2002 to 2003. Prior to this a first HH Survey had been conducted in 1996 to 1997. This permits comparison but also leads to rather typical problems comparing surveys with methodologies differing in details. The data and documents coming with this example are for internal use only. Further information can be obtained form the Mozambican NSI (www.ine.gov.mz). The boxes following are numbered separately and not unusually refer to the Mozambican example.

A first Look at the Story[2]: Box: Poverty Lines in Mozambique

Table : Food and Non-food Poverty Lines

Average exchange rates
(01.01.2002-31.03.2003) [1]   USD   23405

[1] Source: [2] or http:/ www.oanda.com/convert/fxhistory   EURO   24864

		Food Poverty Line	Food Share	Non-Food Poverty Line	Total Poverty Line	USD	EURO
1	Niassa and Cabo Delgado-rural	5434	0.77	1665	7099	0,30	0,29
2	Niassa and Cabo Delgado-urban	7540	0.74	2690	10231	0,44	0,41
3	Nampula-rural	4471	0.75	1501	5972	0,26	0,24
4	Nampula-urban	4853	0.73	1807	6661	0,28	0,27

[2] POVERTY AND WELL-BEING IN MOZAMBIQUE: THE SECOND NATIONAL ASSESSMENT; National Directorate of Planning and Budget, Ministry of Planning and Finance, Economic Research Bureau, Ministry of Planning and Finance, International Food Policy Research Institute, Purdue University (March 2004)

5	Sofala and Zambezia-rural	4155	0.76	1318	5473	0,23	0,22
6	Sofala and Zambezia-urbana	6591	0.75	2183	8775	0,37	0,35
7	Manica and Tete-rural	5629	0.81	1304	6933	0,30	0,28
8	Manica and Tete-urbana	7145	0.74	2545	9690	0,41	0,39
9	Gaza and Inhambane-rural	6614	0.73	2394	9008	0,38	0,36
10	Gaza and Inhambane-urbana	7264	0.68	3457	10721	0,46	0,43
11	Maputo Province-rural	11801	0.70	4963	16764	0,72	0,67
12	Maputo Province-urban	11898	0.65	6398	18296	0,78	0,74
13	Maputo City	12224	0.63	7291	19515	0,83	0,78
	Simple average PL				10395	0,44	0,42

Note: All figures in Meticais per person per day.

1.3.2. *Computing the total poverty line*

This computation involves two steps. The first defines essential non-food basic needs and the second incorporates their cost into the food poverty line (fpl) to arrive the total poverty line (tpl). Simply put, fpl has to be adjusted upward by an amount equal to or proportionate to the cost of procuring the essential non-food basic needs of a poor or nearly poor person. Clearly, essential non-food basic needs require a definition that can be measured. Developing countries generally follow one of three operational definitions or procedures.

1.3.3. *Which are the non-food needs?*

This list is created usually by a group of users and stakeholders in association with the national statistics office or the agency charged with producing the country's official poverty statistics. The list is exhaustive, covering items like clothing and footwear, shelter, fuel and light, household goods, health services, personal care, and education. Costs per person are

assigned to each item. Hence, if nfpl (non-food poverty line) denotes the sum of the costs for these non food items and fpl (the food poverty line) denotes the sum of the costs for these food items the , then tpl the total poverty line: *tpl=fpl+ nfpl*

This was the procedure of choice by some countries during their early years of poverty statistics compilation and some still keep it as part of their official methodology

Simplicity is its main appeal. However, the outcome is very much dependent on a highly subjective list. Adding or subtracting from the list affects tpl directly in an additive fashion. It is easy for anyone to criticize why this item is included while that item is not. Changes in the list would affect the comparability of the tpl time series. Similarly, different lists for different areas or sub-populations result in non-comparable statistics, for example, bread plus rice in cities versus all rice in rural areas or cost of physician attended childbirth in urban regions versus midwife-assisted deliveries in rural areas.

1.3.4. *Share of Food Expenditure and Engel's Coefficient*

A household whose total expenditure per capita (te) is equal to fpl still has to spend for items other than food, and those items must be regarded as essential by the household. The set of non-food items availed of by households for whom te = fpl may then be considered to define essential non-food needs. And the average expenditure of the households for the set may be used to estimate nfpl.

Engel's Law states that household expenditures on food in the aggregate decline as income rise; in other words, the income elasticity of demand for food in the aggregate is less than one a decline toward zero with income growth.

Engel's Law doesn't imply that food spending remains unchanged as income increases: It suggests that consumers increase their expenditures for food products (in % terms) less than their increase in income.

Usually the relation of food expenditure to the total expenditure (fe/te) is called the Engel's coefficient and Engel's Law claims in short, that this coefficient rises more or less? rapidly the poorer the household is.

1.3.5. *Absolute versus relative poverty*

A Poverty line indicates deprivation in an absolute sense, i.e., the value of a set level of resources deemed necessary to maintain a minimal standard of well being. With such a definition, poverty is eliminated once all households command resources equal to or above the poverty line. The $1/day per capita poverty line is one example of an absolute poverty line, but most countries determine their own absolute poverty lines in addition.

Many wealthier countries, on the other hand, set poverty lines based on relative standards. In the EU "relative poverty" is defined as an income below 60 percent of the national median equalized disposable income after social transfers for a comparable household. In Germany for example the official relative poverty line for a single adult person in 2003 was 938 euros per month. For a family of four with two children below 14 years the poverty line was 1970 euros per month.[3] Also bearing in mind the definition is different in terms of absolute / relative poverty and also the distinction between "per person per day" for Mozambique and "single adult person" for Germany has to be made clear, one notices a striking difference in monetary terms.

The relative benchmarks used in Europe reflect the belief that important deprivations are to be judged relative to the well-being of the bulk of society, approximated by the income level

---

[3] On absolute and relative poverty: http://en.wikipedia.org/wiki/Poverty_line and
http://en.wikipedia.org/wiki/Poverty_in_Germany

of the household at the mid-point of the income distribution. In short, inequality matters primarily as a component of deprivation.

1.3.6. *Cost of Basic Needs Approach*

The way in which statistical offices set absolute poverty lines vary but most begin with a cost of basic needs approach.

The basic approach begins with a nutritional threshold chosen to reflect minimal needs for a healthy life, and adjustments are then made for non-food expenses (e.g., housing and clothing). To set a poverty line, statisticians typically identify a basket of foods that will deliver the minimal nutritional requirements. Almost all statistical offices adopting nutritional standards set by the World Health Organization and Food and Agriculture Organization (WHO/FAO). The others set standards based on inputs from national experts.

The calorie requirements vary considerably so let's

Under the cost of basic needs approach, food poverty lines are tied to the notion of basic food needs, which, in turn, are typically anchored to minimum energy requirements. Energy requirements vary depending on age, sex, physical activity levels, body weight, pregnancy status, and breastfeeding status. As the IAF (HH Survey) does not include adequate data on physical activity levels or body weight, we estimated caloric requirements using age and sex as well as estimates

From the 1997 Census on the proportion of women who are pregnant or breastfeeding, Caloric requirements for moderately active individuals by demographic characteristics were obtained from the World Health Organization (WHO 1985). Average per capita requirements in a given region will vary with the average household composition in that region. For example, a region with a greater proportion of children in the population will require fewer calories per capita than a region with a higher proportion of middle-aged adults, as children typically have lower caloric requirements.

In both the 1996–97 IAF and the 2002–03 IAF, the average daily caloric requirement per person per day was approximately 2150 kilocalories in each of the 13 poverty line regions.

To convert the physical quantities of household food consumption in grams to

Kilocalories, a number of different sources were used in decreasing order of preference,

The Mozambique Ministry of Health (Ministério de Saúde 1991)

A food table for Tanzania compiled by the Wageningen Agricultural University (West, Pepping, and Temalilwa 1988)

An East, Central, and Southern Africa food table (West et al. 1987)

The U.S. Department of Agriculture food composition database (USDA 1998)

The U.S. Department of Health, Education, and Welfare food tables (USHEW 1968)

A food composition tables from the University of California at Berkeley

---

look again at the Mozambican example:

Box: Poverty Measuring in Mozambique

The mentioned 2150 kilocalories apply for adults; relative values will be calculated for minors, infants or breastfeeding mothers.

However, neither activity level nor weight is collected in typical household surveys. Thus, while adjustments can be made for age and gender, statisticians must make assumptions about the average activity levels and weights of individuals and different assumptions lead to different nutritional thresholds.

1.3.7. *Households and Individuals: Adult Equivalence*

Poverty is most often seen as an individual property. However, all members of a family may not be equally poor, however. For instance, a grandparent or a child might face deprivation within a household that has adequate resources. To capture this idea, researchers would ideally collect data on individuals, and poverty measurement would take place at the individual level.

In reality researchers collect data on households as collective units (where households are often defined in surveys as those who share meals together or live under the same roof). The simplest way to proceed is to consider the per capita income of the household, calculated by simply dividing total household income by the number of household members. The same method can be applied to total expenditures. This approach is taken, for example, in calculating the widely-used $1/day and $2/day per capita poverty lines.

These per capita calculations weigh all household members identically. A forty-five year-old man is equally weighted as his seventy-five-year-old mother or his ten-year old daughter. And a household with four adults is judged equally poor as another with identical income but with two adults and two young children.

No adjustments are made for cost savings that might benefit larger households relative to smaller ones. The cost of a second child, for example, may not be as great as the cost of the first. And the cost of adding a fourth person to the household often exceeds the cost of adding a fifth. The $1/day approach, though, like many other approaches, fails to account for such changes.

Creating weights that reflect "adult equivalents" helps address the first problem, and adjusting for economies of scale helps respond to the second. The most common approach to establishing adult equivalence standards is to weight, for example, a 45-year old male as "1" and to weight others in proportion to the resources they require. His teenage daughter may take a weight of "0.7" and his elderly mother takes a weight of "0.8". These weightings reflect the fact the daughter and her grandmother consume less than the man to meet their basic needs. In reality, however, it is far from clear how to set specific weights.

Making adjustments for children can matter particularly when comparing changes in poverty over time. If parents give birth to a baby in a given year, per capita income or per capita expenditures will fall substantially for the family since the baby's needs would count as much as anyone else's. But with adjustments that reflect adult equivalence, the addition of the baby to the family - while adding costs - is counted in line with the baby's actual needs.

1.3.8. *Adjustment for non-food needs*

The food poverty line is just one part of the overall poverty threshold. There are two common approaches to making adjustments for non-food needs.

Some countries use the „direct" method (conditional on constructing a poverty line using the „cost of basic needs" approach).

First, necessary items are selected. In the Gambia, for example, the list includes rent, clothing, firewood, transport, education and health costs. In Albania, by contrast, the list also includes tobacco and entertainment. After the list is determined, the goods are priced and the non-food line is formed.

The indirect procedure examines data on food consumption and total expenditures. With a food poverty line in hand, the method entails finding the level of non-food expenditure that would be typical of a household whose food consumption is just at the food poverty line. There are two main ways to do this. The first way is to begin by calculating the „Engel coefficient ", the ratio of food consumption to total expenditures, and then to run a statistical regression that allows prediction of the Engel coefficient for the household whose food expenditure is at the food poverty line. There will be no explication of the regression approach at this place.

A second approach is to calculate the average Engel coefficient for households whose food consumption is in the vicinity of the food poverty line (commonly above or below by 10 percent). In either case, once the appropriate Engel coefficient has been obtained, the overall poverty line can be found by multiplying the food poverty line by the inverse of the Engel coefficient.

1.3.9. *Setting and updating prices*

With the calorie thresholds in place, statisticians can identify a basket of foods that will provide those minimum needs at least cost.

Size and composition of the basket affect the accuracy of the overall poverty line. The trade-off in moving to a larger food basket is mostly given by the added costs of collecting price data. Collecting a moderate-sized food basket (for example with 25 items) but obtaining high-

quality price data will likely enhance accuracy over either smaller or larger food baskets, especially if the latter makes it more difficult to update prices.

Table: Use of prices in setting poverty lines.

	Yes(percent)	Number of observations
Is the general consumer price index (CPI) used to update prices for food basket?	71	34
Is a poverty-specific CPI used?	9	34
Are commodity prices adjusted for differential prices paid by the poor due to lower quantities purchased?	7	30
Are commodity prices adjusted for differential prices paid by the poor due to purchases at different price points?	20	30
Does the composition of the food basket allow for differences in regional consumption habits?	53	32

Source: UNSD international survey of statistical offices, May 2004.

Some adjustment for these concerns is implicit in setting separate poverty lines for urban and rural areas and for different regions. Having multiple lines can add precision, especially in geographically diverse countries, although it drives costs up.

1.4. *International comparisons*

It would be helpful to find a way to slash through the multiplicity of possibilities identified above. One way is to work toward the harmonization of approaches, seeking methodological consensus across countries. Another is to start from scratch with so called international poverty lines. Poverty measures are used both to compare progress across different countries and within a single country. The United Nations and World Bank have adopted $1/day and $2/day per capita poverty lines for international comparisons, even though national poverty lines may be more appropriate for comparisons within a specific country.

The $1 a day poverty line first appeared in the 1990 World Development Report. The World Bank were searching for a universal poverty line, and in their 1990 analysis chose two: $275 and $370 a year at 1985 purchasing power parity prices. This range spanned the poverty lines for 7 low income countries for which recent survey data were readily available: Bangladesh, Egypt, India, Indonesia, Kenya, Morocco and Tanzania. The lower limit of the range coincided with the Indian poverty line. $370 a year quickly became the now familiar $1 a day, while the fact that there had been a lower line was forgotten. Therefore the $1 a day was not constructed: it was a simple arbitrary threshold that could be used to set goals and monitor progress.

As a historical remark: The UN $1/day line happens to roughly approximate India's poverty line in the 1980's. The $1/day line was not constructed, it was chosen as a simple, if arbitrary, threshold that could be used to set goals and monitor progress. One limit of the international lines is that for richer countries, the $1/day line (or even the $2/day line) captures few of those considered poor by experts in the countries themselves. The poverty line in the United States is roughly ten times higher than the $1/day line, for example. The $1/day poverty line used by the United Nations and World Bank is anchored in 1993 international prices, so it is instructive to compare the data to the 1993 poverty line in the United States. In 1993, households in the United States with two adults and two children were deemed poor if their income fell below $14,654 per year or $10.04 per day per person. The per-year figure for the US is thus ten times higher than the international benchmark.

By the way: The Indian government rejected the $1 a day and still does.

Box: About 1 $ per Day

Another problem of international comparison is the translation of US$ into local currency. The simplest approach would be to use official exchange rates. But many goods consumed by the poor are not traded, and official rates can also be distorted by government interventions. It is thus broadly accepted that official exchange rates cannot be relied on to give the appropriate conversion of purchasing power when welfare comparisons are made. These issues have led to a search for alternative methods of conversion.

The United Nations and World Bank use a set of exchange rates calculated as part of the International Comparisons Project. These rates are designed to be used for comparing national income in different countries, and similar to consumer price indices, these rates are calculated based on the relative prices of a set basket of goods in each country. The idea is to calculate conversion factors that aim to equalize the purchasing power of currencies in different countries (which is why they are termed "purchasing power parity" or PPP-- adjustments).

In principle, if a certain kind of man's shirt costs $10 in the United States, then $10 converted via the PPP-adjusted exchange rate should allow a person to have exactly enough money to buy the same shirt in any other country. In practice, the PPP numbers are difficult to calculate, and corrections and refinements to the method continue. A set of numbers is available for a large number of countries, however, and the UN and the World Bank rightly favour them over official exchange rates.

Differences between PPP and official exchange rates are considerable. In late 2003, for example, the ratios of official rates to PPP rates in Brazil, Nigeria, and India were 2.3, 1.4, and 3.3, respectively. This overvaluation of official rates means that the PPP-adjusted figures **raise** the value of $1/day poverty lines when denominated in local currencies and thus the PPP method shows more poverty in the world than would a $1/day line translated into local currencies at official exchange rates.

An alternative approach would be more painstaking but more consistent with the conceptual basis for poverty measurement described above. The idea would be to focus on a set of capabilities that people throughout the world can agree are necessary for living free from the worst deprivations. Elements would likely include having adequate nutrition, health inputs, shelter, and clothing. Each element would be specified carefully in the spirit of the "cost of basic needs" approach described above. The components would be achievable through access to different bundles of goods in different places, recognizing that eating and living patterns vary considerably the world over. The task for statisticians would be to construct locally-relevant commodity baskets that reflect the international consensus on these basic capabilities and to price the baskets using local costs, avoiding the need to use international exchange rates of any sort.

One advantage of this approach is that region-specific poverty lines could be easily accommodated, free of reliance on PPP numbers. The project would no doubt require

considerable international coordination and consultation (unlike current practice), but the reward would be the first truly global poverty approach. Setting new international poverty lines would be a critical first step. As highlighted throughout this course and following [1], coordinating survey techniques and practices would be the second major step toward this ultimate goal. For now, though, many countries will continue to rely upon the $1/day and $2/day per-person poverty lines. They have proved highly effective in focusing attention on world poverty, and they provide a rough benchmark on global trends. But at the same time, their limits should be kept in mind, particularly when completing disaggregated analyses.

1.5. *Harmonization*

A key goal of this course as of the UNSD handbook is to find common ground in approaches to poverty measurement, to better understand differences in approaches, and to sharpen assessment practices. While money-based measures no longer have exclusive hold on our attention, they remain central to analyses. The past two decades of experience, though, reinforce the value of collecting health and education data, as well as other social indicators that describe broader conditions of poverty. Increasingly, researchers also find value in asking about subjective views of poverty and in seeking input on poverty through participatory exercises that involve participants from local communities. Direct measures of access to basic services and infrastructure also provide important inputs in the policy making process. Results from the 2005 United Nations Statistical Division survey on approaches to poverty measurement show a wide range of practices. The diversity partly reflects differences in national conditions and policy needs, but there remain substantial areas where greater uniformity will raise the overall quality of poverty measure and improve comparability of measures across time and location.

The $1/day poverty line approach incorporated into the United Nations Millennium Development Goals demonstrates the power of uniformity, and this introduction points to

ways of going further. Several steps to consider in achieving greater comparability and transparency include decisions to:

• Base poverty measurement on expenditure data rather than income data

• Establish standards for how poverty lines are set, including how to determine both food and non-food portions of poverty lines (where the "cost of basic needs" approach is taken)

• Select a standard set of adjustments for adult equivalence

• Select a standard adjustment for economies of scale

• Create shared guidelines for household survey methods used to collect important consumption items

• Create uniform ways of handling missing data and implausibly low values of income and consumption

• Establish guidelines on whether and how to use data from national income accounts to adjust data from household surveys (many experts will suggest not doing so at all).

These simple steps will bring statistical offices closer toward a common method of measuring poverty. No set of uniform rules and procedures will be superior for everyone all the time, but achieving greater uniformity will be a vast improvement over today's widely varying practices.

1.6. *Exercises*

Study the documents "POVERTY AND WELL-BEING IN MOZAMBIQUE:

THE SECOND NATIONAL ASSESSMENT" ("Poverty and Well Being 2004.pdf") to find out the following for that HHS:

Check the main topics of this chapter in the Mozambican HHS:

How will poverty be measured?

Find out how poverty lines are set. How both food and non-food are portions of poverty lines established?

Module 6: Poverty: measuring and monitoring it for development interventions

Select a standard set of adjustments for adult equivalence (see "Rebuilding after War: Micro-level Determinants of Poverty Reduction in Mozambique")

What kind of adjustments were made to reflect different consumption patterns in Households (Scale economy- see same document).

Analyze the results of Poverty Line Head Counts of different regions and the whole country

2. Poverty Measures

All governments make poverty reduction part of their policy agendas, but how exactly should poverty be measured? This chapter takes up that question with respect to money-based measures: those poverty statistics that measure the degree to which individuals and households fall below a poverty line. Just as there is much diversity in how surveys are collected, the practice of calculating poverty statistics also varies widely.

This chapter focuses on ways in which statisticians aggregate survey data to answer questions such as:

• How many poor people are there in a region?

• How deep is their deprivation?

• Has poverty risen since the last survey?

     and

• Which are other and complementary Poverty Measures?

Any discussion of how to form poverty measures must begin with recognition that statistics have multiple constituencies (e.g., government policy makers, NGOs, researchers, and the general public), and these scattered constituencies often have competing needs and agendas.

Choosing which poverty measure is best depends in large part on the uses to which it will be put. Since no single statistic is likely to answer the needs of all, most statistical offices publish a range of statistics. Recommendations are made for ways to expand the data range to make comparisons easier. **Even better, although not always easy, would be to also make the raw survey data available for others to analyze** (after taking appropriate actions

to protect the confidentiality of surveyed households). That is what we are doing in this course, use raw data to make meaningful statements about poverty.

2.1. Desirable features of poverty measures

Poverty measures are used first and foremost to monitor social and economic conditions and to provide benchmarks of progress or failure. The measures will function well as long as everyone agrees that when poverty numbers rise, conditions have indeed worsened and conversely, when poverty measures fall, that progress has been made.

A second important use for poverty measures is descriptive. Poverty statistics play critical roles in summarizing complex social and economic conditions that inform conversations around economic and social priorities. For this purpose, effective measures need not completely capture all (or even most) morally relevant aspects of poverty. But the limits of measures need to be understood, and transparency and ease of interpretation are critical here.

Let us sum up the desirable features on monetary poverty indicators

Feature	Description
"scale invariance"	Poverty measures should be unchanged if, for example, a population doubles in size while everything else is maintained in the same proportions
„monotonicity"	Holding all else constant, when a poor person's income falls, poverty measures must rise (or at least should not fall).
„transfer"	Holding all else constant, taking money from a poor person and giving it to a less poor person must increase the poverty measure. Conversely, poverty falls when the very poor gain through a transfer from those less poor.
„Transfer sensitivity "	This says that the poverty reduction in the case in which a very poor person is made better off relative to her neighbour should be greater than the reduction in the case in which the recipient is less poor.
„decomposability"	The poverty measure should be decomposable by sub-population. Sub-populations may include, for example, residents of different regions. sub-groups have to be distinct from each other (so that there is no overlap in membership) and that together they encompass the entire Population.
single, unique measure that	One candidate that satisfies all of the axioms on which we agree to identify

Feature	Description
would be fully "characterized" and encompass all the prior axioms	"poverty" Unfortunately: This candidate is impossible to find, seen the previous discussions

Box: Desirable features of monetary poverty indicators

2.2. A First Example: The Headcount measure

The simple headcount index is the most used poverty measure, but it violates several important axioms. Of the four measures described below, the one that satisfies most of the desirable axioms above--the Watts measure--turns out to be the least used. These two facts suggest an ongoing tension between the desire for simplicity and transparency pitched against the desire for rigor. The measures below will be compared in that light.

The headcount is the simplest and best known poverty measure. It identifies the share of a population whose income/ expenditure is less than the poverty line. It is, not surprisingly, the most commonly calculated poverty measure. The measure literally counts heads, allowing policymakers and researchers to track the most immediate dimension of the human scale of poverty.

The headcount is calculated by comparing the income/ expenditure y_i of each household to the poverty line z. (The index i = 1..N, where N is the total number of households in the sample.) Concretely, an indicator variable is constructed for each household, taking the value 1 when income/ expenditure falls below the poverty line or 0 if income/ expenditure is greater: $I(y, z) = 1$ if $y_i \leq z$

$I(y, z) = 0$ if $y_i > z$

> In the easiest form the number of poor in the sample is divided by the total number of the sample. This number, usually in form of a percentage, is called the headcount measure.

If a household is poor is found out by comparing the Expenditure by person per day with the Poverty Line. Is the expenditure below the PL, then the household is considered to poor, otherwise it is considered not to be poor.

2.3. Poverty Indices in Detail

This chapter treats details of measures of poverty. Since we started with monetary indices, the main focus will be on these.

2.3.1. Foster-Greer-Thorbecke (FGT) Indices

The most widely used index is the Foster-Greer-Thorbecke (FGT) index

$$P_\alpha = \frac{1}{Nz^\alpha} \sum_{i=1}^{I} (z - y_i)^\alpha$$

, where j is a sub-group of individuals with income / expenditure below the poverty line (z), N is the total number of individuals in the sample, y_i is the income / expenditure of individual i and α is a parameter that allows us to distinguish between the alternative FGT indices. When α is equal to 0, the expression simplifies to I/N or the headcount ratio, a measure of the incidence of poverty.

Poverty depth is measured by the poverty gap, which is obtained with α equal to 1. The severity of poverty (or squared poverty gap) is measured by setting α equal to 2

Module 6: Poverty: measuring and monitoring it for development interventions

In a more specialized case, when we want to compare Poverty Indices of sub groups, this formula varies a little. When it comes to sample weights (roughly a number standing for how many units of the universe are represented y one sampled item), this formula becomes even more complicated.

The un-normalized simple Foster-Greer-Thorbecke poverty index FGT P(k; z; α) for the population subgroup k is as follows:

$$P_{k;z;\alpha} = \frac{1}{\sum\limits_{i=1}^{n} sw_i^k} \sum_{i=1}^{n} (z - y_i)^{\alpha}$$

The normalized FGT index, which is normally used, is defined by:

$$\overline{P}_{k;z;\alpha} = \overline{P}_{k;z;\alpha} / z^{\alpha}$$

i	Observation number
Yi	the value of the variable of interest (e.g. expenditure) for observation i
w	the Sampling Weight.
wi	the Sampling Weight for observation i.
s	the size variable.
si	the size of observation i (for example the size of household i or the equivalent adult value).
swi	wi* si
k	the group variable (e.g. Age group, Sex of Head of HH, Urban/Rural).
ki	the group of observation i.
c	The selected group variable
swik	swik=swi if ki = c, and swki=0 otherwise.
n	Sample size

2.3.2. More in Headcount measure

Module 6: Poverty: measuring and monitoring it for development interventions

The headcount index is simply the sample average of the variable I(y, z), weighted by the number of people in each household ni. The measure is calculated by first counting the number of poor individuals I weighted by number of people in the household, divided by all the HH in the sample: I/N.

In reality, not the poor individual is counted but the poor household and then weighted (divided by Number of household members or equivalent adults).

When the sample is not representative of the underlying population (e.g., if the sampling strategy involved random stratification), population weights should also be included in the calculation (see remarks on Sample Design and Survey comparison for further discussion). The sample weight has nothing in common with the weight mentioned in the previous paragraph. While the first takes account of the size and composition of the household, the latter accounts for the different sizes of selected groups, especially in multi-stage sampling. This weight adapts the sample outcome according to the size of the expected sub-sample. It adjusts for different sub-sample sizes which are intrinsic to the sample design. As an example: the sub-sample size for an urban area covers less people than for a rural area, so the sample area for, e.g. "SA 1" stands for 104 households in the population (this in simple terms is the reciprocal of its probability of selection into the sample). For another, a rural area this probability (or sample weight), e.g. "SA 2" will have a value of 530. It is not in itself a difference between rural or urban, it tells something about a probability of household to be selected into the sample, in this case this probability is much higher in the first than in the second sample area [35] I agree, this has to be corrected –see above

The headcount is an important descriptive tool. As a sole guide to allocating resources,
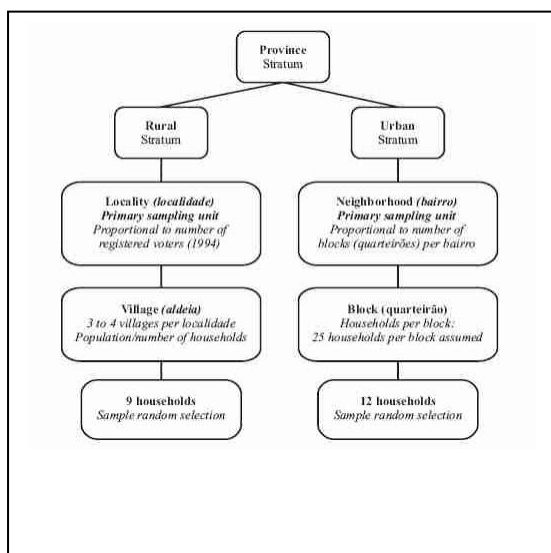
though, the headcount can significantly mislead. There are two large tensions. First, the

headcount registers no change when a very poor person becomes less poor. Nor does the

headcount change when a poor person becomes even poorer. One can argue though that

changes in the income / expenditure distribution below the poverty line matter (only) in a

moral sense. This notion is captured by the transfer axiom above, but the headcount fails the

test.

Graph: The Sample Design for Mozambique Household Survey 02/03

A second tension flows from the failure of the transfer axiom, combined with the focus on

So if efforts are allocated specifically to reduce the headcount, priority will likely go to helping the least poor over helping the poorest.

whether people are above or below the poverty line. If policymakers see their task as

reducing poverty as measured by the headcount, their work will be made easier by focusing

on improving the lot of individuals just below the poverty line. A little improvement at this level

can raise the incomes / expenditures of the "barely poor" below the poverty line and hence

can reduce the poverty headcount fairly rapidly. Directing resources to very poor people, on

the other hand, may be socially beneficial, but far larger income gains are required to take

them over the poverty line and thus to make a dent in the poverty headcount.

The headcount remains a highly valuable measure, even if, when used on its own, it is a

poor guide for resource allocation. One step to make the approach more useful is to

calculate the headcount for "sub-poverty" lines at lower thresholds than the overall poverty

line. These may capture, for example, the income required to purchase the food basket only,

excluding non-food needs. In this context it is important to mention the „decomposability"of the FGT indices, allowing to calculate for subgroups.

2.3.3. Poverty gap

This second widely-used measure has a problem similar to the headcount: it is descriptively very useful but, if used alone, would also serve as a poor guide to resource allocation. The poverty gap measures the amount of money by which each individual falls below the poverty line. It matters here whether income / expenditure and the poverty line are measured on a per capita basis or whether they have been put into adult equivalent terms or adjusted for scale economies. The appropriate formulas are given below. The starting point is to calculate the total shortfall in income / expenditure for the poor population:

$$P_{z;1} = \sum_{i=1}^{M} (z - y_i) I(z, y_i) n_i$$

where the poverty line is z, income / expenditure is y, I(z,yi) is a 0/1 indicator of poverty for each household, household size is ni, the total number of households in the sample is M, and individuals are indexed by i. The calculation gives the total sum of money that would be needed to make up for the gap between the existing incomes / expenditures of the poor and the official poverty line. The calculation above is correct only if income / expenditure is in per capita terms. When income / expenditure is made instead in adult equivalent terms (or adjusted for scale economies), the correct calculation is:

$$P_{z;1} = \sum_{i=1}^{M} (z - y_i) I(z, y_i) a_i$$

where ai gives the number of adult equivalent units in household i. The figure above, also called the shortfall may be helpful for budget planners, but it obscures the sense of individual

deprivations. An alternative is to instead calculate the average shortfall for the population below the poverty line:

$$P_{z;1} / I(\text{Number of Poor})$$

When viewed together with the headcount, this version of the poverty gap measure shows the distance (on average) to be travelled in raising incomes to leave poverty behind. Because the figure is denominated in currency, conversion to a common international currency (e.g., Euro or Dollar) will aid global comparisons.

A different approach that can enhance comparability is to divide the index by the poverty line:

$$P_{z;1} / (I(\text{Number of Poor}) * z)$$

Normalization puts the average gap in terms of the percentage shortfall from the poverty threshold, freeing the measure from denomination in a particular currency. The measure is now easily comparable across countries and across time, a helpful improvement. Routinely publishing poverty lines alongside the normalized poverty gap and the headcount would allow observers to calculate for themselves all three of the poverty gap variations described above.

Dividing the index by N, the total population of the sample, combines the three data points (headcount, poverty gap, and poverty line) to form another widely-used variant of the poverty gap.

2.3.4. Squared poverty gap

One way to transform the poverty gap described above into a distributionally sensitive measure is to raise the individual gaps to a power greater than 1. Foster, Greer and Thorbecke propose a class of measures built on this idea which have found their way into much of the poverty analysis published by the World Bank. With income/ expenditure expressed in per capita terms, the measures take the form:

$$P_{z;2} = \frac{1}{N} \sum_{i=1}^{M} \left(\frac{(z - y_i)}{z} \right)^2 I(z, y_i) a_i$$

When income / expenditure is in household terms, the adult equivalent size a_i should be replaced with the household size variable n_i.

For a FGT with α > 1, the measure is distributionally sensitive and the particular case in which α = 2 (often referred to as the squared poverty gap) is now the most widely-used distributionally-sensitive measure.

By squaring the poverty gap, improvements in the resources of the poorest individuals count most, since they are the ones for whom the initial resource gap is largest. The measure satisfies the transfer axiom but not the transfer-sensitivity axiom. To satisfy the latter, the poverty gap would have to be raised to a higher power, cubed rather than squared, say. Cubing adds .transfer sensitivity, a property that many find appealing. But it puts very heavy weight on the well-being of the poorest. Perhaps weight that would be judged too great in a social calculus at levels of α between 1 and 2, not only is transfer sensitivity not satisfied but the reverse holds: holding all else the same, a regressive transfer among the very poor increases poverty *less* than a same-sized regressive transfer among the moderately poor, a clearly undesirable feature. Distributional-sensitivity is achieved by weighing deprivations of the poor inversely to their base incomes. There are many ways to do this, and the weighting

scheme in the squared poverty gap has the advantage of relative simplicity. The simplicity can help provide some intuition in understanding why the poverty measure moves over time.

2.3.5. **Watts index**

A simple poverty measure that satisfies the transfer axiom was first put forward by Watts (1968), who argued for the following measure:

$$WIndex = \frac{1}{N} \sum_{i=1}^{M} \left(\ln(z) - \ln(y_i) \right) I(z, y_i) a_i$$

as above, when income / expenditure is in household terms, the adult equivalent size *ai* should be replaced with the household size variable *n i.*

The measure is distributionally-sensitive by virtue of its use of logarithms. The character of the logarithm function means that the Watts index is much more sensitive to changes in the lowest incomes than it is to changes for those with higher incomes. That is, transferring $10 to a very poor person counts as a far larger contribution to poverty reduction than transferring $10 to a richer (but still poor) neighbour.

Allocating anti-poverty resources to minimize the Watts index would thus tilt efforts toward the poorest, which is a feature that some analysts find appealing. The Watts index also satisfies the transfer-sensitivity axiom described above, and it is decomposable into the population-weighted sum of the poverty indices of regions or groups. (The squared poverty gap shares this feature too.) Being decomposable is useful when a population can be divided into a number of distinct groups or regions. Poverty measures can then be calculated for each group or region, and, if the poverty measure is decomposable, the individual poverty measures can be aggregated (using population shares as weights) to form the overall poverty measure for the entire population. Decomposing poverty measures in this way can help to pinpoint the groups or regions contributing most and least to overall poverty.

2.3.6. Comparing the measures

Obviously all the different indices mean different things

The headcount numbers the poor (in relation to the total population)

The poverty gap shows the relative distance (on average) the poor have to travel in raising incomes to leave poverty behind.

The squared poverty gap shows the distribution of poverty among the poor, the higher it is the bigger are the differences among the poor

The Watts Index measures in log-terms the differences between the poor and the poverty lines

2.4.   Exit time and the value of descriptive tools

For all of the theoretical appeal of the distributionally-sensitive measures described immediately above, the headcount remains by far the most common poverty measure in use. The Millennium Development Goals, for example, focus on reducing the headcount of poverty below $1/day, rather than minimizing a distributionally-sensitive measure. One reason for the continuing use (and usefulness) of the headcount is its descriptive properties. It is a simple means for illustrating the scale of poverty. In this sense, it is an intrinsically meaningful measure. The poverty gap is also intrinsically meaningful, taking us from counting people to counting shortfalls of income / expenditure or consumption. It answers the question: how much would have to be spent to eliminate poverty through costless (and perfectly) targeted transfers.   Its underlying assumptions are clearly unrealistic: in practice, transfers will never be administratively costless, nor will they ever be perfectly targeted. However, this hypothetical question still provides a helpful way to quickly gauge the scale of deprivation.

The "average exit time" is based on a similar sort of hypothetical question. The underlying assumption is as unrealistic as that under the poverty gap, but the measure nevertheless can fruitfully frame discussions of poverty. The measure is based around the number of years that it would take poor households to grow out of poverty given a hypothetical, steady growth of income / expenditure. In practice, income / expenditure growth will seldom be steady over long periods, nor will all poor households be able to grow at the same rate. But, as with the poverty gap, asking the hypothetical question provides a quick way to gauge one important aspect of the condition of poverty. Hypothetical exit times are simple to calculate. If the assumed growth rate of income / expenditure is g percent per year, an individual whose income / expenditure starts at yi will take T years to exit, where T solves this equation:

$$z = y_i(1 + g)^\tau$$

The equation can be solved by taking logarithms, yielding that the number of periods of growth required before exit is T = ln(z/yi) / g. Of course, T = 0 for all households already above the poverty line.

The only data required are the poverty line z, the assumed growth rate g, and the median income / expenditure of the population below the poverty line yMean. So If the hypothetical exit time, Ti, for each poor household is averaged over the population below the poverty line, the "average exit time" A is:

$$A = \frac{\ln(z / y_{Mean})}{g}$$

So If the hypothetical exit time, Ti, for each poor household is averaged over the population below the poverty line, the "average exit time" A is:

$$A = \frac{1}{N} \sum_{i=1}^{I} \frac{\ln(z/y_i)}{g} I(z, y_i) n_i = \frac{1}{N} \sum_{i=1}^{I} T_i n_i$$

This equation is analogous to the average poverty gap described in appropriate equation, and it shares similar weaknesses and strengths. Its chief strength is its simplicity and descriptive value. Its main weakness is that when a less poor household exits poverty and all else is unchanged, the average exit time, A, will fall. This makes the exit time not a very good candidate to be the sole measure of poverty. However, it can still be a very useful component for the analysts. As the FGT Indices the average exit time is that it can be decomposed explicitly to show the impact of inequality below the poverty line.

The exit time has a useful relationship to another established measure. If exit times are calculated for the entire population of a country (with those above the poverty line having 0 exit times), the average turns out to be simply the Watts poverty index divided by g, the hypothetical annual growth rate

This measure, the "population average exit time" naturally shares all properties of the Watts index, satisfying both the transfer axiom and the transfer-sensitivity axiom. But it has the addition of a new interpretation, akin to the interpretation of the poverty gap described by the corresponding equation.

With economic growth very much a part of the poverty reduction policy agenda, tools like exit times provide ways to summarize data in a manner relevant to policy debates on growth-based poverty strategies. They complement the other measures described above.

Importantly, it should be remembered that exit times describe possibilities based on simple assumptions as used here, the exit times are not based on actual forecasts or careful predictions. These simple exit times, though, can be useful in identifying opportunities and constraints to guide policy.

2.5.   Measurement Limits and Errors

No survey is perfect, but some collection methods are far more reliable than others. Particular problems arise when expenditures (or incomes) are either substantially over-counted or under-counted, and the biases can be exacerbated by the choice of poverty measure. Under-counting leads to exaggerations of poverty rates, and the distributionally-sensitive measures described here are particularly susceptible to the exaggeration of under-counting in the lower tail of the income distribution (Squared Poverty Gap and Watts Index).

2.5.1. Mean

The result of summing the ratios and dividing the result by the total number ratios.

e.g. . the mean of daily consumption per EQAdult in a province

2.5.2. Median

The value such that number of ratios less than this value and the number of ratios greater than this value are the same.

e.g. the median of daily consumption per EQAdult in a province

2.5.3. Quantiles / Percentiles

Quantiles are points taken at regular intervals from the cumulative distribution function of a random variable.

There are several special cases of Quantiles: Percentile is the value below which the specified percentage of cases falls. So the median value is the 50 Percentile.

Commonly used Quantiles are Quartiles(25,50,75), Quintiles (20,40,60, 80) and Deciles (10,20,30,40,50,60,70,80,90), the Percentile in its original meaning (1,2,…98,99)

2.5.4. Lorenz Curve

The Lorenz curve is a graphical representation of the cumulative distribution function of a probability distribution; it is a graph showing the proportion of the distribution assumed by the bottom y% of the values. It is often used to represent income distribution, where it shows for the bottom x% of households, what percentage y% of the total income they have. The percentage of households is plotted on the x-axis, the percentage of income on the y-axis. It can also be used to show distribution of assets. In such use, many economists consider it to be a measure of social inequality.

Graph: Lorenz Curve for Expenditure in Sofala
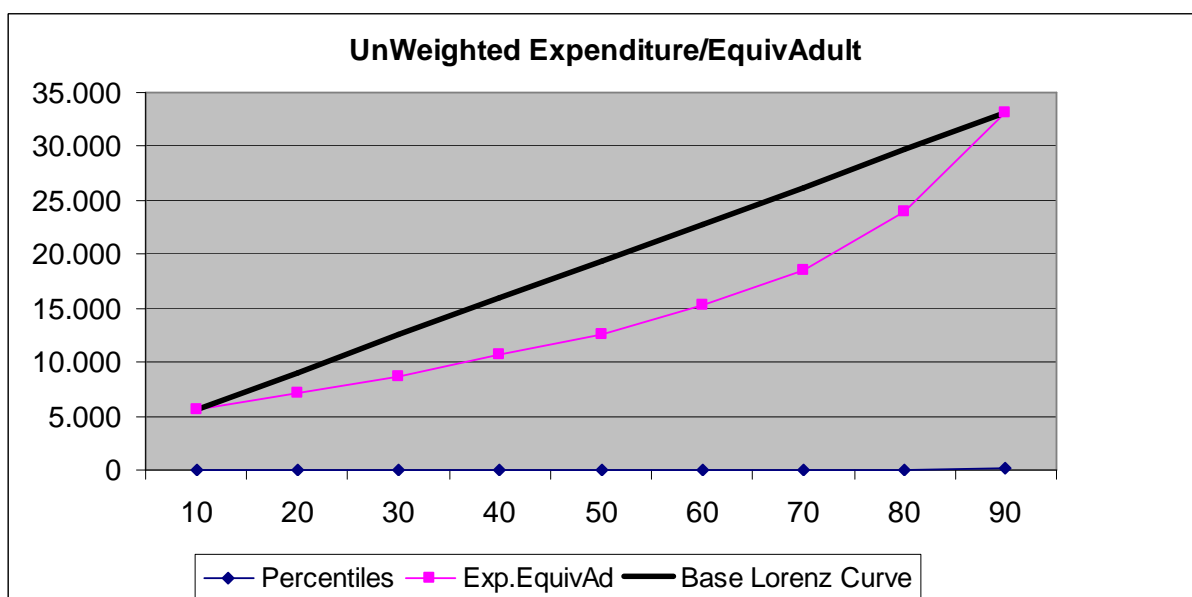
Source: Graph from "HH Survey Sofala"

If all individuals are the same size, the Lorenz curve is a straight diagonal line, called the line of equality. If there is any inequality in size, then the Lorenz curve falls below the line of equality. The total amount of inequality can be summarized by the Gini coefficient (also called the Gini ratio), which is the ratio between the area enclosed by the line of equality and the Lorenz curve, and the total triangular area under the line of equality. The degree of asymmetry around the axis of symmetry is measured by the so-called Lorenz asymmetry coefficient.

The Lorenz curve shows in graphical form the distribution of Household expenditures or income broken down by Quantiles. For the exercise the weighted expenditures were split into 10 equidistant groups. The Lorenz curve would be straight at a 45º inclination, if the expenditure would be equally distributed. The result for Sofala show that the reality is far from this.

It is somewhat simpler to look at the Gini-Coefficient to have numerical indicator for the equality or inequality of e variable distribution. The Gini- Coefficient (between 0 an 1) measures the area between the Lorenz curve and the Base Line, it would be 0 for a complete fit (all households groups have the same share of the income/expenditure) and 1 for extreme inequality (like one household has all, the rest has none). This means the closer the Gini coefficient is to 0, the more uniform the distribution of expenditure is. The Gini Coeffient is the ratio of the mean of expenditure among the group members (the 10 percentiles) divided by the group mean.

2.5.5. Gini Coefficient

$$Gini = \frac{1}{(2n^2 y_{Mean})} \sum_{i=1}^{n} \sum_{j=1}^{n} |y_i - y_j|$$

The Gini coefficient (or Gini ratio) is a summary statistic of

the Lorenz curve and a measure of inequality in a population. The Gini coefficient is most easily calculated from unordered size data as the "relative mean difference", i.e., the mean of the difference between every possible pair of individuals, divided by the mean size.

If the y values are first placed in ascending order, such that each y has rank i, the some of the comparisons above can be avoided and computation is quicker and easier: - where y is an observed value, n is the number of values observed and i is the rank of values in ascending order.

$$G = \frac{2}{n^2 \, y_{Mean}} \sum_{i=1}^{n} i(y_i - y_{Mean})$$

The Gini coefficient ranges from a minimum value of zero, when all individuals are equal, to a theoretical maximum of one in an infinite population in which every individual except one has a size of zero. It has been shown that the sample Gini coefficients defined above need to be multiplied by in order to become unbiased estimators for the population coefficients.

2.6.   Concluding Remarks

Focusing on the most commonly used measures promotes comparability across countries. These include the headcount index, poverty gap, and squared poverty gap. This section has described how to calculate and interpret these measures (in addition to the Watts index), and identifies their respective strengths and weaknesses.

Statistical offices can go further, though, by also publishing simple statistics that provide a richer picture of conditions. These statistics are seldom very costly to compute and can substantially enrich analysis. The first is the median income of the poor population. The

median gives the income level below which the bottom 50 percent of the poor population lives. This simple measure indicates whether the bulk of the poor population is close to or far away from the poverty line. In section 3.5 above, it was also shown how median income can be employed in the exit-time framework.

Going further, it would be helpful to also publish the income of households at the 25th percentile and the 75th percentile of the income distribution below the poverty line. Ultimately, publishing the entire Lorenz curve (the mapping of population shares to income/expenditure shares) would be most revealing and would add little extra cost. The median, though, is the natural place to start, followed by incomes at other important focal points of the distribution.

One of the most valuable steps that statistical offices can take is to put in place ways to make the raw data for poverty analyses available to researchers. Steps would have to be taken to secure confidentiality to households in the survey, but fortunately methods to do so are now well-established. Broadening access to data will allow analysts to better compare conditions and to develop new tools that can ultimately benefit statistical offices, policy makers, and citizens.

2.7. *Exercises*

Compare the mentioned indices in a simple spreadsheet exercise

(PovertyLineExerciseEx3_1.xls).:

What happens to the Poverty Indices, if all the expenditure of households is raised by 100

CU

What happens, if the expenditure of the two "least poor" households is raised by 100 CU

What happens, if the expenditure of the two "poorest" households is raised by 100 CU

What happens, If money is transferred (100 CU each) from all the "not poor" households to

"poor households"

What happens, If money is transferred (100 CU each) among the two "least poor"

households to two "poorest" households

Please note the results of the exercises in a table comparing the 8 Indices, like

Exercise	unweighted				weighted				Comment
	P_0	P_1	P_2	Watts	P_0	P_1	P_2	Watts	
1									
2									
3									
4									
5									

Please calculate the average exit time for Sofala and the population groups "Urban" and

"Rural" separately. Calculate for two hypothetical growth rates 3% and 10%. Use prepared

Exercise sheets (PovertyLineExerciseEx3_2.xls)

	Type of Index	Total	Urban	Rural
6	Average Exit Time (3%)			
7	Average Exit Time (10%)			

Watch the calculation and the way EXCEL is used to calculate the Gini Coefficient for the Total Population of Sofala and calculate Gini Coefficients for Urban and Rural Population separately

	Type of Index	Total	Urban	Rural
8	Gini			

3. Presenting Data useful for Poverty Reduction Strategies, Plans and Action

This section focuses on the formulation of poverty reduction policies. It shows how various poverty tools can be of considerable value to policy makers in strengthening the poverty alleviation impact of government spending. Poverty profiles can play an important role in understanding poverty and formulating poverty reduction policies. In this chapter, we provide some country specific examples to illustrate how poverty profiles can be constructed and how they can be utilized to design policies.

The primary step in determining the degree of poverty is establishing a poverty line that specifies in monetary terms a society's judgment regarding the minimum standard of living to which everybody should be entitled[4]. Once the poverty line is determined, one can construct poverty profiles, which provide overall estimates of poverty, distribution of poverty across sectors, geographical regions and socioeconomic groups, and a comparison of key characteristics of the poor versus non-poor.

Geographical targeting is also becoming an important means for channelling public resources to the poor. Many governments use it to target programs, such as food aid, public works, and delivery of health care and education. This approach is commonly referred to as

---

[4] Watch the difference in definition to the previous chapters

"poverty mapping". This chapter provides a brief review of methodology used in the construction of poverty maps. It also points out the effectiveness and limitations of poverty mapping.

3.1.   *Poverty monitoring and poverty profiles*

The three poverty indices discussed in the previous section are often used as a tool to monitor poverty over time at the aggregate level. Needless to say that monitoring poverty at the aggregate level is important because policy makers want to know if the government policies are helping the poor.

Poverty profiles show how poverty varies by geography and subgroups across society. Divisions include regions, communities, sector of employment, and household size and composition. Profiles can also show how rates of economic growth in different sectors and regions affect aggregate poverty. Accordingly, poverty profiles are extremely useful in formulating the most effective economic and social policies to combat poverty. They identify regional location, employment, age, gender and other characteristics of the poor. This information can be used to formulate poverty alleviation policies. Profiles can also help answer a wide range of questions such as:

Who are the poor?

Where do they live?

What do they do?

On what sectors do they depend for their livelihood?

Do they have access to economic infrastructure and support services such as social services and safety nets? ,

How can the government target resources to them?

Look at the document "Poverty and Well Being 2004.pdf" [2] and consider to answer the questions above

The three FGT poverty measures -incidence of poverty, poverty gap ratio and severity of poverty- have the attractive property of being additively decomposable poverty measures (see previous section). This property is essential and quite useful in analyzing poverty profiles.

Contribution of each region to total poverty can be used as an indicator for allocating public assistance to each region. Since most of the poverty is found (where?), government spending to reduce poverty should be concentrated in that region. There is no consensus about the distribution of economic growth across various socioeconomic and demographic groups. Household survey data can be used to investigate how economic growth affects poverty among various groups.

Although poverty profiles are very useful in understanding the nature of poverty, they are limited to showing bivariate associations between various socioeconomic groups and poverty measures. In other words, they do not control for other omitted variables, which also have an impact on poverty. In many instances, this profiling approach can generate misdirected policies.

This poses the otherwise crucial question: "Which is the reason for poverty?"

We may construct poverty profiles by simple model transformation, regressing the probability of being poor on a large number of relevant household characteristics that are generally used in poverty profiles. From these models, one can estimate the marginal effects, or elasticity, of probability of being poor with respect to any explanatory variable included in the model. We will do some exercises on this in the next chapter. The main attraction of these models is that

we can isolate the effect of a single variable by controlling for all other variables included in the model.

To illustrate this point, consider the document "IFPRI_hasDevelopmentbeenProPoor.pdf" [8] to discuss the point that the wealth has increased among the poor. The IFPRI document reflects the views of the official politics. Another document illustrates quite a different picture "Joseph Hanlon_Poverty.pdf" [9] using a different approach.

3.2.   *Capability deprivation*

The income-poverty line, which identifies the poor from the non-poor, can never perfectly distinguish between individuals who are able and unable to enjoy a minimum set of capabilities Thus, it is important to investigate whether the poor suffer greater capability deprivation than the non-poor. If they do, more effective policies can be devised to raise their living standards, such as providing cash or in-kind transfers or greater access to government services. This section investigates whether the poor (defined in income terms) actually suffer greater capability deprivation. Looking at the following table:

Table 12: Changes in school enrollment, 1996--2002

Percentage of children 7 to 17 years old (at the time of the survey) who:	IAF96	QUIBB00	IAF02
Ever attended school			
National	60.8	69.0	79.7
Rural	54.8	62.1	75.0
Urban	82.6	82.3	89.4
Were attending school (at the time of the survey)			
National	48.8	61.3	67.8
Rural	43.7	55.0	62.5
Urban	67.7	73.5	78.8

Table: Comparing Scholl Enrolment in Sofala

This table does not indicate any differences between poor and non-poor as it could do to show the difference, although the QUIBB00 would not provide for any poverty figures. However it shows the differences in enrolment between urban and rural households. What political decision will be the result of these findings is another issue altogether. The distribution of wealth or diminution of poverty is the next challenge. The result of not alleviating these differences: excluding social or regional strata from economical welfare will heighten tension and increase social unrest.

Africa with different communication standards might react a bit different from other world region, but with record cell phone increase rates for sub Saharan Africa, the communication gap is closing for the urban population, at least in countries with governments open to the western economical system.

This raises the further question: "Which are the indicators of capability deprivation to be observed?"

Indicators of capability deprivation can be:

Drinking Water

This index measures the quality of drinking water as it is measure in the MDG Report: the proportion of population using an improved drinking water source

Toilet or Sanitary Facilities

Human waste disposal is another important factor related to people's capability to live a healthy life. Unhygienic toilet facilities can spread infectious diseases. The index of toilet

facilities measures quality of toilets available to a household or as in Human Development Reports: HDR theProportion of population using an improved sanitation facility

Cooking Fuel

Gas and electricity are the cleanest and most convenient fuels for cooking. But they can be expensive, and they may not even be available in the areas where poor people live. The index of cooking fuel reflects its cleanliness and convenience.   It also reflects health status, particularly for women who are using solid fuel in an enclosed space.

Availability of Electricity

Percentage of the population with access to electricity

Housing Condition

Data on the number of rooms (and The number of sleeping rooms) in each dwelling. The data are used to calculate the rooms (and sleeping room) available per 100 persons or as MDG also recommends number of people room (overcrowding is >3).

Access to household consumer durables

Data on the use of bicycles, telephones, air conditions, and washing machines

Productive assets held by poor and non-poor

One of the important reasons why poverty persists is that the poor do not possess productive assets (Fixed assets, Financial assets, Housing etc.) and the productivity of the assets they do own may be low

Many developing countries use micro credit to help the poor acquire productive assets. There are many other policy options, such as marketing training to help the poor get better prices for their produce and services. However, the more challenging issue is devising

policies that would be targeted to the poor. Poverty mapping that helps identify the poor is an increasingly important tool to better target anti-poverty programs

3.3. Poverty mapping

Geographic targeting is becoming an important tool for allocating public resources to the poor. It is commonly regarded as a more efficient way to reduce poverty than untargeted, universal programs. Many governments in developing countries are giving greater importance to decentralization, whereby the district or provincial governments play an important role in poverty reduction policies. To implement such policies, it is important to know the spatial distribution of poverty. Poverty mapping is the spatial analysis of poverty. It maps the incidence of poverty within each region and sub-region of a given country. A number of methods have been devised to measure spatial distribution of poverty. There is not enough space in this chapter to present all the methods that have been used in practice.

Household surveys are the most important data source for measuring poverty. However, usually their sample sizes are too small to provide precise estimates of poverty for small geographical units, such as provinces and districts. Alternative data sources are population censuses, which do not suffer from small sample problems, but typically provide very limited information from each household. For instance, censuses do not offer information on households, consumption expenditures or incomes, preventing income poverty from being measured directly. However, small-area estimation is a statistical technique that combines household survey and census data to estimate income poverty at small geographical units. It has been used by the U.S. government for planning and targeting. And recently, the World Bank staff have refined this technique and applied.

Module 6: Poverty: measuring and monitoring it for development interventions

The techniques of poverty mapping will not be discussed in detail. The main obstacle for most African countries is the data availability on the most basic level. This will be discussed in the next chapter. More detailed information in [1, Chapter VII]

Construction of poverty maps requires having access to census data at the household level. Statistical offices of many countries do not allow, for reasons of confidentiality, such detailed information be made available to individual researchers. Some statistical authorities, however, make available aggregated census data, which unfortunately, leads to loss of precision of poverty map estimates, particularly at the lower level of disaggregation. A further requirement of poverty mapping is that the household surveys have a large subset of variables that are also used in the census, which may not always be the case. However Household Surveys could be adapted to increase the harmonization with the census. If this is not the case variables that are available in both household survey and census may not be sufficiently correlated with the household consumption. In this case, the regression model will not be able to predict poverty maps accurately. Finally, poverty mapping assumes that the explanatory variables (e.g. the poverty indicator: consumption) in the household survey are produced from the same data-generating process as the census data. This assumption, however, can be statistically tested. The minimum requirement for this assumption to hold is that both household and census surveys should correspond to the same period. The maximum allowable time difference will depend on the rate of economic change that is taking place in the country. Many countries do not have census and household surveys for the same period.

In Mozambique the HH Surveys were conducted in 1996/7, 2002/3 and 2008/9, the censuses in 1997 and 2007.
Look at the Recommendation of IMF's GDDS concerning poverty data
http://dsbb.imf.org/Applications/web/gdds/gddstableb/

Box: Proposing Poverty Measures by IMF's GDDS

Core indicators	Encouraged extension(s)	Periodicity	Timelines
Income poverty: number and proportion of people or households with less than minimum standard of income or consumption; valuation of minimum consumption bundle	Measures of the distribution of household or per capita income or consumption	3-5 years	6-12 months following the survey
Other poverty measures: measures of deprivation or insecurity used to identify the population living in poverty, such as evidence of malnutrition, endemic diseases, educational achievement, and lack of access to basic services	Separate poverty estimates for urban and rural populations or for major regions, states, or provinces. Disaggregation of data by sub national or regional units, as appropriate		

In most developing counties, the census is conducted every ten years. Household surveys, however, are conducted more frequently. The ten-year period is too infrequent, leading to the creation of poverty maps that are outdated long before the next poverty mapping exercise is undertaken. Outdated poverty maps can lead to misallocation of scarce public resources. Given so many problems in combining household survey and census, an alternative method of constructing partial poverty maps is proposed later. This approach does not require the use of census data.

For the purpose of formulating a poverty reduction policy, one wants to know which districts are poor so that policymakers can target policies to them. The first task is to define a poor district. Since the poverty rate at the national level was 38.6 percent in 1997-98, it is reasonable to assume a district to be poor if more than 50 percent of its population is living in poverty. The null hypothesis is that the percentage of poor people in a district is 50 percent or less. The alternative hypothesis will obviously involve districts where more than 50 percent of the population is poor. Thus, one can identify a district as poor if one rejects the null hypothesis at the 5 percent significance level.

$$(1) \qquad 100 * \sqrt{\frac{0.5 * 0.5}{n_i}}$$

$$(2) \qquad p_i > 50 + 1.67 * 100 * \sqrt{\frac{0.5 * 0.5}{n_i}}$$

National Poverty Index
Percentage

Percentagem
- 36,1 - 44,6
- 44,7 - 53,6
- 53,7 - 63,2
- 63,3 - 80,7
- Missing Value

Source
MOZ_MF, Pobreza e Bem - Estar em Moçambique Segunda Avaliação Nacional, 2004
MOZ_MF, Understanding poverty and well being in Mozambique 1996 to 1997, 1998

Note: The boundaries and the names shown and the designations used on these maps do not imply official endorsement or acceptance by the United Nations.

If pi is an estimate of the percentage of poor in the ith district based on a sample of size ni, then its standard error under the null hypothesis will be (1) Using a one-tailed test, the hypothesis will be rejected at the 5 percent significance level if (2):

3.3.1. Poverty Mapping and National Comparison with DevInfo

DevInfo 5.0 (or later) is an advanced database management system to keep track of the commitments towards human development. It is endorsed by the UN to report on the progress towards the Millennium Goals. With DevInfo, you are "literally a few mouse clicks away from facts that help you to" (☹ this is PR speak for DevInfo but anyway):

• Analyze data for evidence-based decision making.

• Undertake results-based monitoring and evaluation.

• Link different levels—national, sub national and regional—of planning.

Standard Poverty Mapping with DevInfo for Mozambique
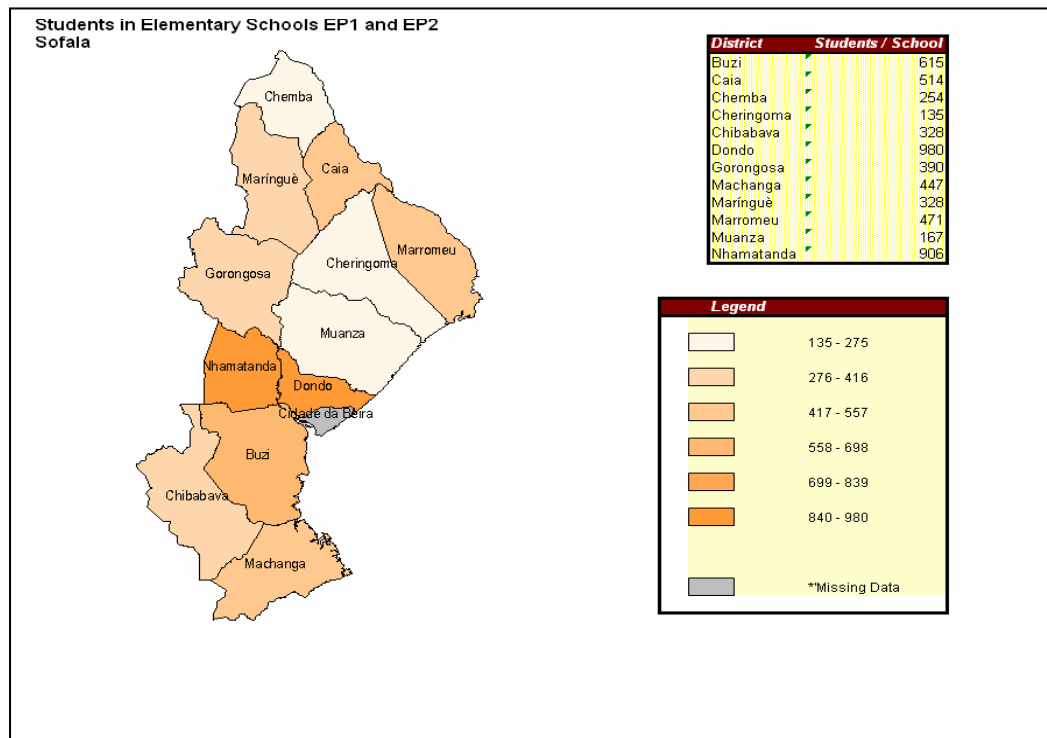
Customized Poverty Mapping with DevInfo for Sofala

The advantage of DevInfo Data is, that the owner has full control over the quality and source of the data contrary to sources like WDI (World Development Indicators). The DB Development Pack (DevInfo <Current Version> Data Admin) is readily available for download and allows for display of customized data. It goes without saying that GIS Data (coordinates for the poverty maps) have either to be available or have to be substituted or added by the Data Admin Pack.

Module 6: Poverty: measuring and monitoring it for development interventions

Typical household surveys, which are the basis of most poverty estimates, comprise a sample of several thousand households. Although this provides a rich information base on the living standards of the sample households, the sample size is usually only sufficiently large to estimate poverty to the first subnational administrative unit, such as a province or state.

By combining the detailed information of a household survey with the comprehensive coverage of a national census, one may estimate poverty levels for much smaller areas. Although these small area estimates are indirect and are calculated with a certain degree of statistical error (or uncertainty), they may be suitably precise to be useful for policy purposes ([16])

At least two data sets are required to implement the method. One is a detailed household survey that includes a measure of welfare, which is typically consumption per capita. The other data source is a national census or, alternatively, a large national survey that includes a significant share of the country's population.

In the first stage, the analysts use multiple regression analysis to estimate a model of household consumption. The variables in the model are restricted to those variables that are available in both the survey and the census, and the data sources are examined to ensure that they are indeed comparable. See a good exercise on regression

In the second stage, the estimated model parameters are applied to the census data. Simulation methods are used to introduce a random disturbance term, which is required because the model does not predict consumption perfectly. The simulations provide an estimate of consumption per capita for every household in the census. These estimates are then used, along with the appropriate poverty lines, to estimate poverty measures at various levels of aggregation, such as province, state, district, subdistrict, and municipality. The method also produces an estimate of the standard error of the poverty measure, which is used to construct a confidence interval for the poverty estimate. The estimates are then typically merged with a map in a geographic information system (GIS) to facilitate the presentation and visual analysis of patterns.

The World Bank has since developed special-purpose poverty mapping software, **PovMap**, that is freely available at http://iresearch.worldbank.org. **PovMap** (http://iresearch.worldbank.org/povmap/PovMap2/PovMap2Main.asp) has a graphical user interface that greatly simplifies the small area estimation of poverty and inequality.

Also **PovcalNet** a comparison program of the WB for comparing Poverty Estimators can be accessed on the same site (http://iresearch.worldbank.org/PovcalNet/jsp/CChoiceControl.jsp?WDI_Year=2007)

What is said below about accuracy and transparency is true here. Nevertheless the programs have their value for international comparison.

Box: Methods for combining HH surveys and Census data for using small area estimates

3.3.2.International Comparison with WDI (World Development Indicators)

Poverty gap at $1.25 a day (PPP) (%) 0.4% 40.8%
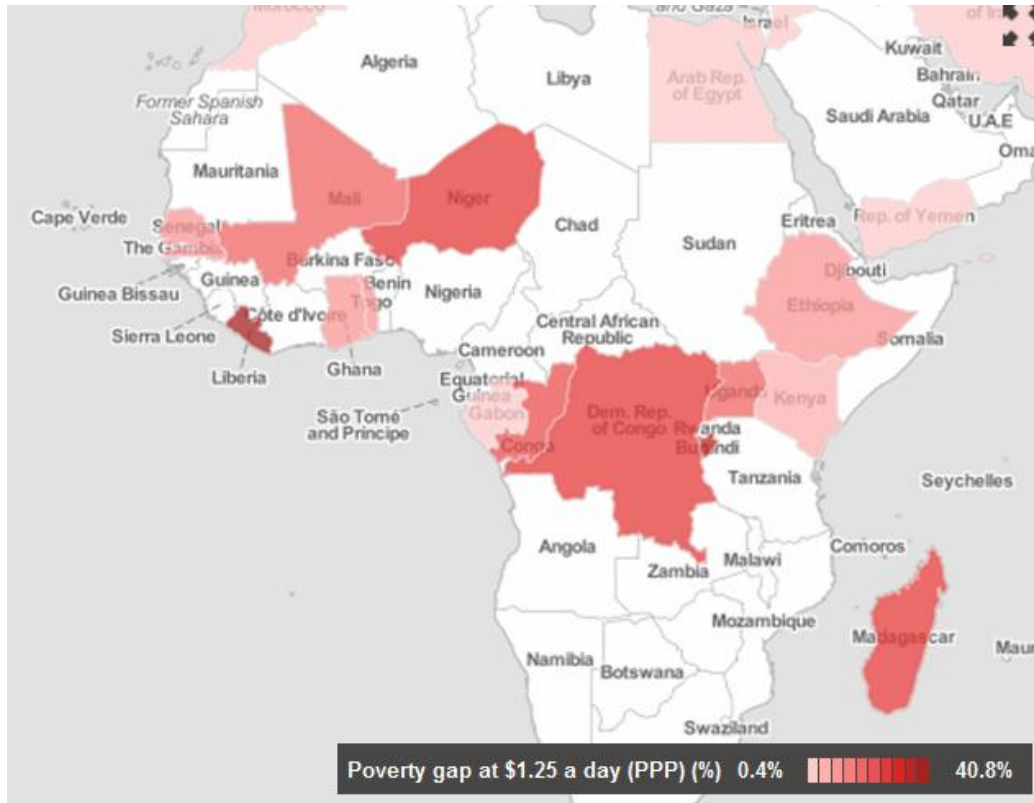
Standard

Poverty Mapping with WDI for Africa (Link [3])

These Data are nice for general comparison but are not suited for in-depth comparison. Moreover the origin of data is not always transparent.

3.4.   Special Subjects towards a Poverty Monitoring System

This section deals with additional information suitable for adding value to the formulation of poverty reduction policies. It shows how additional and cross-sectional techniques can enhance the various poverty tools. It is unrealistic to think that international standards will

enable full comparison of all countries and regions. The standardization of statistics about industrialized countries is rather advanced (see OECD statistics) and an economical necessity. Rarely can these standards be transferred easily to countries on different levels of economical development. It is of fundamental importance to adjust appropriate means for sound political measures in these countries. Without this and the permanent support of the international community, the capacity of national leadership cannot succeed to improve local well-being and reduce poverty.

3.4.1. Non-income measurement methods

The minimum basic needs (MBN) (also called unmet basic needs (UBN)) approach has been used in a number of countries instead of or in addition to the income-based basic needs approach. In this approach non-monetary indicators representing different dimensions of poverty are chosen, estimated and monitored. The subset of Millennium Development Goals [MDGs] minus the income indicators provides excellent

Examples:

Proportion of underweight children to represent malnutrition

School enrolment, primary school completion, and youth illiteracy rates to represent basic education;

Infant and under-five mortality rates, maternal mortality ratio, and births attended by skilled health staff to represent primary health care; and the

Ratio of girls to boys in primary and secondary school, ratio of literate females to males, proportion of seats held by women in parliament, and

share of women employed in the non-agriculture sector to represent the gender equality dimension or goal.

Note the parallelism to some MDG Indicators

Module 6: Poverty: measuring and monitoring it for development interventions

Many, though not all, of these indicators are long-term outcomes or output indicators. Case in point: a child being underweight is the result of years of chronic under nourishment. Also, these UBN indicators are expressed in different units of measure. This has made producing a composite index a difficult and perpetually subjective task. However, this has not prevented agencies, particularly international bodies, from constructing such indexes. These include the Human Development Index (HDI - http://en.wikipedia.org/wiki/Human_Development_Index)[5] and other indices that UNDP puts out annually for each country in *Human Development Report (*http://hdr.undp.org/en/reports*).* These may have added value more as devices for advocacy than as monitoring tools. It shows clearly the limits of international comparison. Few developing countries compile composite indexes, preferring to use the indicators individually and collectively in much the same way that they will be used to monitor progress in the MDGs.

National statistical systems have also been compiling many of the UBN indicators long before the international development agencies declared poverty reduction their overriding strategic objective. Many are extracted from population and housing censuses, demographic and health surveys, civil registries and other administrative reporting systems. In fact, choice of indicators in a country's UBN information system is often determined by existing data collection systems; seldom is a new system established just to meet the additional requirements for new indicators.

However, unlike income poverty statistics compilations, there are differences in the selection of dimensions and indicators for the basic needs, partly owing to variations in data availability. The three broad categories of basic needs often considered are dwelling

---

[5] The HDI – human development index – is a summary composite index that measures a country's average achievements in three basic aspects of human development: health, knowledge, and a decent standard of living. Health is measured by life expectancy at birth; knowledge is measured by a combination of the adult literacy rate and the combined primary, secondary, and tertiary gross enrolment ratio; and standard of living by GDP per capita (PPP US$).

characteristics, access to safe water, and access to sanitation facilities. Basic education and economic capacity (e.g., GDP growth rate) are sometimes included in an expanded UBN set of indicators. In the ECLAC, the UBN approach has a solid conceptual foundation as it measures actual satisfaction or dissatisfaction of needs rather than the capacity to satisfy them. In this light, it is complementary to the income poverty line approach. Dimensions of basic needs chosen are often those highly correlated with income, so much so that they have been used to identify households under extreme poverty. Assessment and monitoring of poverty through the UBN approach is far from widespread in Africa.

The main poverty dimensions considered are basic education, primary health, and housing characteristics, such as access to safe water, toilet facilities and building materials used. UBN methods can and are being brought down to sub-national levels.

When it comes to updating poverty indices, the frequency of HH surveys, the main source of income / expenditure measures is not sufficient to reflect changes or long term trends. Other measurement systems have to be considered to complete a national or regional poverty monitoring system like DHS (Demographic and Health Survey) or MICS (Multiple Indicator Cluster Survey) as sources of data for non-income measures.

3.4.2. Poverty Dynamics

The previous text focused almost exclusively on analysis of poverty at a single point in time. Yet, in a given time period, people may be poor either because they have always been poor or because they have suffered a negative shock that temporarily pushed them below the poverty line. With a single cross-sectional survey, it is difficult to separate these two types of poverty even though each may require different policy prescriptions.

Therefore, with the ambition to recommend some guidelines for poverty monitoring, the following text extends the analysis to many time periods, and thus is concerned with the dynamics of poverty.

Module 6: Poverty: measuring and monitoring it for development interventions

Examining changes in poverty over time raises difficult issues. But it also provides a richer and more realistic portrait of the nature of poverty. Individuals and households typically live for many decades, which implies that a person's poverty status may change over time. If it does not change over time, it would be trivial to extend static analysis to dynamic settings. As will be seen below, the poverty status of many individuals and households appears to change a great deal over time, a finding that is surprising to both researchers and policymakers. This chapter assumes that income / expenditure is an effective variable for measuring welfare. While this assumption may narrow the scope of poverty analysis, it is needed to keep the scope of topics manageable.

Even with a single-variable study, many important issues can arise in dynamic analysis that are not simple to resolve. Thus, despite the increased interest in poverty dynamics collecting and analyzing survey data on poverty dynamics is a difficult task for any statistical agency. The chapter starts by examining three important conceptual issues in poverty analysis:

Relationship between income inequality and poverty at a single point in time and income mobility over time,
Distinction between chronic and transient poverty
Issues concerning the measurement of income growth among the poor.

The possibility that people's poverty status can change over time raises several conceptual issues. This section discusses three of the most important:

Understanding the relationship between income inequality and income mobility at a single point in time (which has direct implications for the relationship between income mobility and the dynamics of poverty),
Distinguishing between chronic (long-run) and transient (short-run) poverty
Measuring income growth of the poor.

3.4.3. Defining Poverty

Now we could try to compare the mentioned FGT indices with other models and measures. It is an example to apply statistical techniques like cluster analysis and factor analysis to existing survey data. It should encourage statisticians to do their analysis on their own terms. It goes without saying that the results have to abide with the confidentiality of statistics and have to be supervised and commented prior to publication. This example will only be referred to in the bibliography [4] and [5].

It will be the purpose to define a base line for development rather than poverty to compare future results in developing the province. It goes without discussion that reducing poverty is one of the foremost aims of projects, planning and politics, but economic poverty/consumption is a rather limited figure among others. It is rather volatile, might change quickly due to natural and economic casualties. It has moreover some methodological drawbacks as mentioned in other literature. One important shortcoming is the de-facto exclusion of extra-vulnerable target groups. It is clear, that these groups, for example homeless, street-children etc. do not constitute the important group at which development is aimed, but their existence and their socio-demographic development constitutes an important indicator of society's processes. There has been extensive discussion of this topic in the document about measuring development on district level based on a qualitative survey carried out in Sofala Another aim would be to identify predictors, parameters most likely to reduce poverty. At last the expenditure distribution is compared among the different districts. Where are the most striking inequalities to be detected and where can one observe a rather uniform pattern of expenditure (poverty) [7]

A display of result together with a graphical representation of the overall SMD [6]index is

shown below

District Development in Sofala as seen by the System of Monitoring Development SMD:

Results 2005/2006/2007

---

[6] It is proposed to use a qualitative questionnaire, designed as in [7]. The questionnaire is very simple can be

answered by the head of household or the substitute and would not required neither extended length of

permanence in the household nor extended processing requirements. It contains basically 18 questions with 5

possible answers (usually from *deteriorated very much*.. to… *improved very much*) This technique has been

applied to longitudinal surveys, monitoring development over time and permits to get quick responses with quick

tendency results. The qualitative results are elaborated and indexed allowing immediate information about

improvement ore deterioration in the questioned area. The result of these interviews is of course opinions and not

facts. Therefore the quality of results have to be controlled by the next quantitative survey, preferably the Census,

IAF again or the QUIBB. By using a survey over a longer period, both quantitative and qualitative surveys can

complete each other

Legend:
- Negativo
- Positivo
- Sem Dados

Map regions: Chemba, Caia, Marínguè, Cheringoma, Marromeu, Gorongosa, Muanza, Nhamatanda, Dondo, Cidade da Beira, Buzi, Chibabava, Machanga

Indices -SMD	2005	2006	2007*
Buzi	-14%	-11%	
Caia	-16%	-9%	
Chemba	7%	-8%	
Cheringoma	0%	-7%	40%
Chibabava	-4%	-21%	
Dondo	-8%	-14%	
Gorongosa	19%	17%	43%
Machanga	0%	-11%	
Maringue	-1%	-12%	24%
Marromeu	8%	12%	
Muanza	2%	-1%	35%
Nhamatanda		-31%	
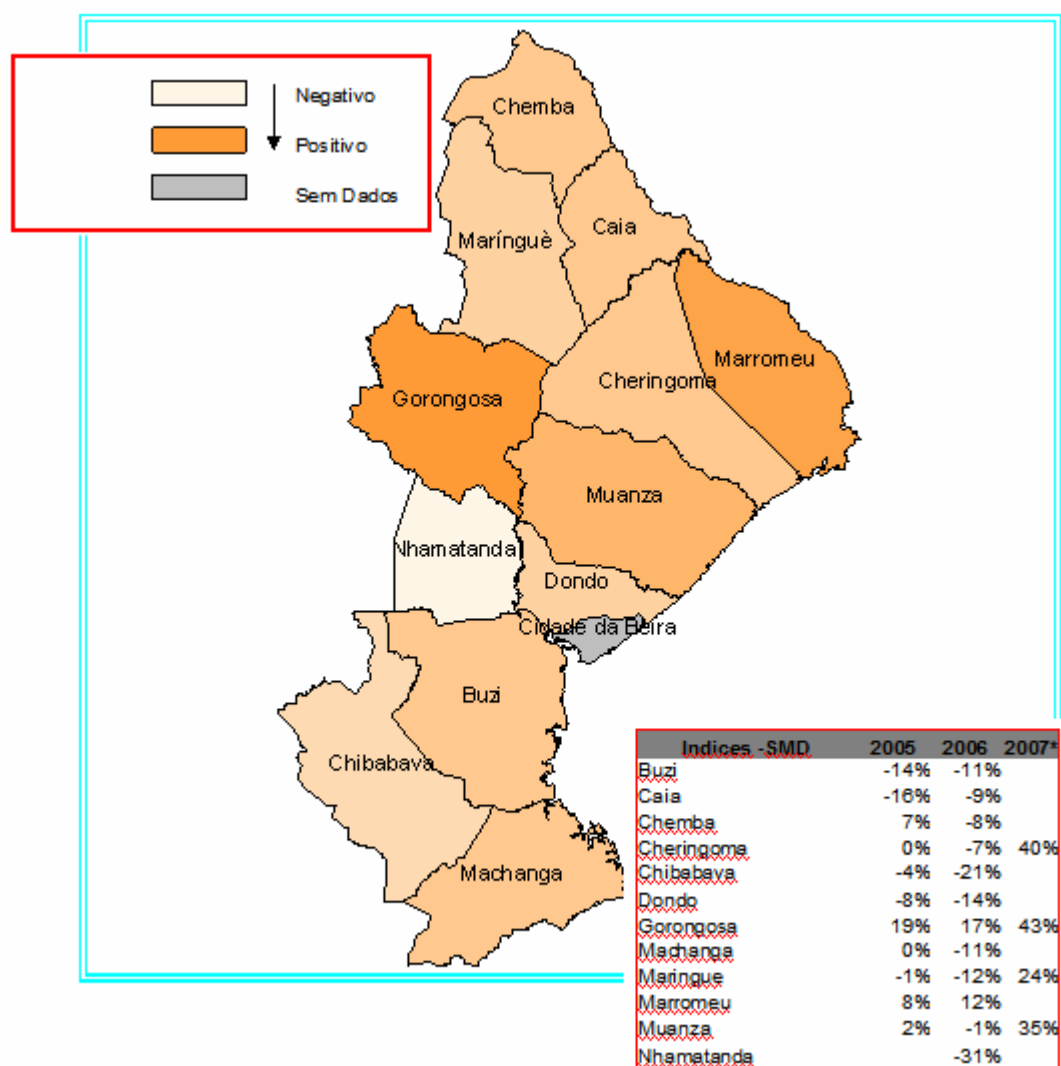
3.5. *Exercises*

Use the evaluation document: UN- QUESTIONNAIRE ON POVERTY MEASUREMENT from the UN Handbook (378 pp)

Support document: [1]

Complete it in groups. It is not important to complete the whole questionnaire (use only part A: poverty line)

Recommend alongside these recommendations a set of indicators for a political measure to be able to combat poverty. (Use poster "Poster_RecomendationsPovertyMeasures.doc")

Think of as poor area/ poor section of society

How are changes measured until today

Name indicators used so far

How could these be improved, what could be done better

Which indicators would be needed

Which indicators have higher or lower priority

Which are your strategies to communicate poverty figures

Try to estimate costs for a 10 year period for the statistical Surveys or other indicator proposed

Consider a Poverty Reduction Strategy based on your consideration

Bibliography

[1] HANDBOOK ON POVERTY STATISTICS_UN_2005.pdf

[2] Poverty and Well Being 2004.pdf

[3] WB-Using Census and HH Data for Poverty (link [6])

[4] IFPRI-Using Census and HH Data for Poverty Analysis: IFPR-Census use- Poverty, Inequality, and Geographic Targetingfcndp192.pdf

[5] Klaus Röder-Project Proposal for a Survey Measuring Development in Sofala on District Level: ProposingSMD_Sofala070301.pdf

[6] IFPRI_hasDevelopmentbeenProPoor.pdf

[7] Joseph Hanlon_Poverty.pdf

Links

[8] http://data.worldbank.org/indicator/SI.POV.GAPS/countries?display=map

[9] http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTPA/0,,contentMDK:21094736~isCURL:Y~menuPK:435761~pagePK:148956~piPK:216618~theSitePK:430367,00.html

Better Monitoring for Better Evaluation

Module 7: Case studies (2): choosing indicators for effective monitoring

Contents

1. Introduction and aims

The aim of this module is to demonstrate through practical examples how indicators suitable for monitoring may be chosen with reference to the logical framework.

In module 4 we studied two development strategies, one for basic education in The Gambia and the other for health reform in Nigeria. The logframes which we developed will be the starting point for the work of the present module.

The task is to produce a limited set of indicators suitable for monitoring progress at the Goal, Purpose and Output levels, following the recommendations of having not more than two indicators for Goal and Purpose, and not more than three for each output. We shall also provide indicators for the super-goal, if appropriate. Where possible, we shall indicate potential sources of information for each indicator, though this will of necessity be a little academic since we do not have perfect information about the data sources for the countries in question.

The final table, therefore, will be of this nature:

	Narrative	Indicator Source	Comments
Super-goal			
Indicator 1			
Indicator 2			
Goal			
Indicator 1			
Indicator 2			
Purpose			
Indicator 1			
Indicator 2			
Output 1			

Indicator 1.1			
Indicator 1.2			
Indicator 1.3			
Output 2			
Indicator 2.1			
Indicator 2.2			
Indicator 2.3			
Output 3			
Indicator 3.1			
Indicator 3.2			
Indicator 3.3			
Output 4			
Indicator 4.1			
Indicator 4.2			
Indicator 4.3			

We shall start with the Education Sector Medium Term Plan for The Gambia. This document lists some summary indicators, but the list does not cover the full range of outputs of the strategy. We shall therefore attempt to develop a reasonable set of education-related indicators which measure the quality of policies and standards, as well as educational outcomes.

Our second example is the Health Sector Reform Programme of the Federal Republic of Nigeria. Although it is complex – we saw in Module 4 how a nested series of logframes would best describe the totality of the strategy – the central section of the document does contain a large number of indicators. We shall consider these and their suitability for monitoring, and try to choose two or three which would be the most appropriate for each level of the logframe.

The Nigerian example also provides logframes at different levels of the hierarchy – for the high level strategy and for work in specific sectors – which allow us to demonstrate a broad range of indicator types.

For both examples, the text of the relevant strategy will provide the starting point for the development of indicators. The full texts are to be found in the references.

Note that the description of an indicator should not be confused with the target. For example, "infant mortality drops by 5 percentage points from 2003 to 2006" is not an indicator. The indicator is "infant mortality rate" and the target is "5 percentage point drop between 2003 and 2006".

For this exercise our concern is with finding a suitable indicator. To complete the exercise, it will be necessary to establish a baseline and targets for each indicator, as we described in Module 5.

2. Republic of The Gambia Education Sector Medium Term Plan 2008-2011

2.1. *Introduction*

The Plan document is Reference A. To summarise, The Gambia has a 10-year Education Sector Strategic Plan (ESSP) which is itself derived from VISION 2020. An evaluation of progress towards the education targets showed that there were a number of specific areas which required additional effort. The principal ones which were identified during the exercise to construct the logframe in Module 4 are:

- Quality: improvements are required in both teaching and teaching materials;
- Access: classrooms and equipment are inadequate and children are unable to get access to basic education in some areas;
- Equity: the balance between trying to get more girls into school while retaining the boys is not satisfactory – more work is needed on gender policies and standards;
- Management: the entire education system needs better management to deliver the services which are required.

The logframe we proposed in Module 4 is:

Super-Goal	Reduction of poverty in The Gambia
Goal	To create a human resource base in The Gambia with the education, training and skills required by VISION 2020
Purpose	To deliver relevant and high-quality free basic education to all Gambians
Output 1	Better quality teaching and teaching materials
Output 2	Comprehensive provision of classrooms and equipment in rural and urban areas
Output 3	Effective use of improved standards and policies for gender-sensitive basic education
Output 4	Improved management of the education system

2.2. *General comments on data*

Data on the education sector in The Gambia are not readily available on the web. The website of the Gambia Bureau of Statistics (GBoS) refers to two datasets, GAMINFO (based on the international DEVINFO database) and GNADA, the Gambia National Data Archive. The latter is a repository of census and survey reports. The former has education data for selected years and selected indicators, but there are many gaps.

The website of the UNESCO Institute for Statistics (UIS) gives a good measure of what education data have been produced and quality controlled by UIS for adherence to norms, standards and international definitions. The database is reasonably complete on basic statistics such as gross and net enrolment figures and rates, for both primary and secondary schooling, but there is little on tertiary education. Much of the data on expenditure is estimated. There is good coverage of numbers of teachers.

The DHS surveys have not been done in The Gambia.

The Gambia participated in the MICS surveys in 2000 (MICS2) and 2005-6 (MICS3) which provide certain key estimates including pre-school attendance, school readiness, net intake, net primary and secondary attendance, primary completion and young adult literacy.

The website of the DoSBSE, www.edugambia.gm, contains no data more recent than 2007 and has no access to EMIS. A Google search for EMIS turns up reference to a plan to renovate EMIS during 2006/7. Two World Bank contracts for updating EMIS in 2009 and 2011 are also advertised on the web. The implication is that EMIS, if operating at all, is not operating effectively and the data could be unreliable.

From this external view, then, we cannot say with any certainty what may and what may not be available. For the exercise, we shall therefore concentrate on what our ideal indicators would be.

2.3. *Purpose indicators*

The Purpose of the MTP is "to deliver relevant and high-quality free basic education to all Gambians". The outcome will be children who are better educated. Table 1 on page 18 of the MTP gives the key indicators:

- Intake rate
- Enrolment (gross and net)
- Completion rate
- Gender parity
- Literacy
- Pupil-teacher ratio
- % teachers qualified
- Annual instruction hours
- Pupil : textbook/instructional materials ratio

For the purpose statement we are looking for two indicators which between them capture the essence of the programme. It seems that the key messages of the MTP are a) that all children should come to school and b) that they should remain there and receive a good quality education. The UIS website gives us some useful hints in its Glossary:

Gross intake ratio to the first grade of primary – it indicates the general level of access to primary education. It also indicates the capacity of the education system to provide access to grade 1 for the official school-entrance age population

Net enrolment rate – it indicates the extent of coverage in a given level of education of children and youths belonging to the official age group corresponding to the given level of education

Gross enrolment ratio – it shows the general level of participation in a given level of education. It indicates the capacity of the education system to enrol students of a particular

age group. It can also be a complementary indicator to net enrolment rate by indicating the extent of over-aged and under-aged enrolment

For definition of completion, we turn to the MDG handbook.

Primary completion rate – this monitors education system coverage and student progression. It also measures human capital formation and school system quality and efficiency

It would seem then that the two best indicators for our Purpose statement are intake and completion.

1.   Gross intake ratio to the first grade of lower basic school;
2.   Completion rate of upper basic school.

2.4. *Indicators for Output 1 – better quality teaching and teaching material*

On page 37 and 38 of the MTP the critical issues which impact on teacher quality are identified. These include teacher absenteeism, poor or outdated pedagogical knowledge, few professional development opportunities and limited or non-existent pedagogic support. The proposals to deal with these problems are less clear and seem (page 38) to involve mainly improving the monitoring and supervision of the situation. Later in this chapter there are references to teacher training, in relation to specific issues such as disability, and one mention of the introduction of a Higher Teachers Certificate (Primary).

Pages 24 and 25 also refer to improving teaching quality for basic education but again without a plan of action. One statement says there will be a strategy for "training and retention of teachers in the system" but does not say how they are to be trained and retained.

Page 13 notes that strategies for improving learning outcomes will include stepping up the training and supply of teachers, curriculum review and reform, and provision of relevant instructional materials.

Bringing these things together, it would seem that

a) The system needs more teachers;
b) They need to be better trained;
c) They need motivating to attend school;
d) They need more support;
e) They need proper teaching materials.

We could therefore propose three indicators as follows:

1.1   Number of newly trained (and re-trained) primary teachers entering the system;

1.2   Teacher absenteeism rate (or achieved annual instruction hours);

1.3   Pupil : Textbook/instructional material ratio.

2.5. *Indicators for Output 2 – comprehensive provision of classrooms and equipment in rural and urban areas*

Pages 24 and 25 have a number of activities specific to this output. These include:

- Creating a school environment conducive for learning
  - Fencing
  - Drinking water
  - Toilets
- Increasing the learning opportunities in basic education
  - Classroom rehabilitation
  - Furniture
  - School lunches
  - Transport from remote areas
  - Opening new feeder schools where numbers permit

Section 3.2, pages 41 to 44, refers specifically to plans for constructing classrooms and has targets for this based on the expected increase in intake.

Three indicators for this output could then be:

2.1   Number of classrooms built or rehabilitated and equipped to meet standards;

2.2   Percentage and number of primary schools with inadequate facilities (particularly for health and hygiene) in rural and urban areas;

2.3   Number of children being taught in primary schools with inadequate facilities.

2.6.   *Indicators for Output 3 – effective use of improved standards and policies for gender-sensitive basic education*

Gender mainstreaming is a cross-cutting issue in the MTP, described in section 3.3, pages 44 to 46. The main Education Sector Strategy was girl-friendly, but this unfortunately led to a drop in enrolment for boys. The current policy is to be child-friendly, ensuring equity and gender parity. In other words, all children should go to school and there should be nothing in the education system which is unfair to either sex, whether this be a lack of toilets for girls, or a lack of male teachers, or a curriculum biased towards boys, or sexual harassment, to name just a few.

Statistical data should be disaggregated by sex – male or female – wherever possible. Gender analysis is a more complex issue.

Gender analysis is the collection and analysis of sex-disaggregated information. Men and women both perform different roles. This leads to women and men having different experience, knowledge, talents and needs. Gender analysis explores these differences so

policies, programmes and projects can identify and meet the different needs of men and women. (Source: UIS Glossary)

Both the Purpose indicators can give information for boys and girls separately. We could also modify indicator 2.2 to record in addition schools which specifically have inadequate facilities for girls, and indicator 2.3 to record the numbers of boys and girls separately.

Other education indicators should be equitable, for example net enrolment, progression to secondary education, performance in maths and science, and literacy.

We can therefore suggest three indicators for this output as follows:

3.1   Literacy rates among 15-24 year olds, by sex;

3.2   Performance in maths and science in the National Assessment Test, by sex;

3.3   Number of qualified primary school teachers, by sex.

2.7.   *Indicators for Output 4 – improved management of the education system*

Chapter 4 of the MTP (pages 59 to 61) briefly describes the problems with the current management structure but says little about how exactly to put it right. Service Level Agreements are clearly important (page 60) and have specified targets which will need to be reflected by the monitoring framework. But the remainder of the chapter describes systems and policies – such as "effective financial planning and management ensured" – without very much more information.

For the system to be managed effectively and efficiently, as it says at the top of page 60, people and departments need to be clear about their responsibilities and need good

performance management systems. We can also add that the departments need to be correctly staffed with properly trained people. Finally, the people need to operate the systems in a way which is responsive to changing needs and situations. The SWOT analysis on page 86 also reminds us that management includes governance, right down to the level of school boards.

In the absence of detail, suitable indicators are hard to define. In reality, the process would be one involving all stakeholders. For the purpose of this exercise, and drawing to some extent on the indicators developed for Output 2 of the Nigerian example, we shall propose the following:

4.1   Number of education-related strategies and planning frameworks developed and judged fit for purpose;

4.2   Review and assessment of roles and responsibilities within the education system;

4.3   Service Level Agreements: percentage of sectoral objectives achieved in time.

2.8.   *Goal indicators*

The Goal is to create a human resource base in The Gambia with the education, training and skills required by VISION 2020.

We now turn to the goal of the MTP, which is shared with the wider ESSP and VISION 2020. As it says on page 12 of the MTP, the ESSP is grounded in the overall goal of ensuring that by 2015 universal access to relevant and high quality education has been achieved. The expectation is that this will enable all Gambians to lift themselves out of poverty and become a "well-educated, trained, skilled, healthy, self-reliant and enterprising" people.

Our super-goal will contain poverty indicators, and our purpose statement contains intake to and completion of basic education. Our goal indicators should therefore be situated between these. We could for example refer to the ESSP and overall education-related outcomes (refer also to chapter 2 on post-basic education); or we could move on to the way a better education is used, such as gaining employment and creating income. Perhaps a combination of the two would be suitable.

1.   Completion rate of secondary education, by sex;

2.   Unemployment rate, by sex.

2.9. *Super-goal indicators*

The super-goal is the reduction of poverty in The Gambia.

We propose to select two indicators from the MDG list in order to monitor poverty.

1.   Proportion of population below national poverty line (or $1 a day at 1985 ppp prices);

2.   Share of the poorest quintile in national consumption.

2.10. *Summary*

A monitoring framework for The Gambia Education Sector Medium Term Plan now looks like this. Given the lack of information on data sources, that column has not been completed. Note the modifications to the Purpose indicators, and indicators 2.2 and 2.3, to monitor the gender outcomes.

MTP	Narrative	Indicator Source	Comments
Super-goal	Reduction of poverty in The Gambia		

Indicator 1	Proportion of population below national poverty line (or $1 a day at 1985 ppp prices)		MDG indicator
Indicator 2	Share of the poorest quintile in national consumption		MDG indicator
Goal	To create a human resource base in The Gambia with the education, training and skills required by VISION 2020		
Indicator 1	Completion rate of secondary education, by sex		Continuation to secondary is also a measure of the quality of primary education
Indicator 2	Unemployment rate, by sex		This would be expected to decline if the ESSP and VISION 2020 are working
Purpose	To deliver relevant and high-quality free basic education to all Gambians		
Indicator 1	Gross intake ratio to the first grade of lower basic school, by sex		
Indicator 2	Completion rate of upper basic school, by sex		MDG indicator
Output 1	Better quality teaching and teaching materials		
Indicator 1.1	Number of newly trained (and re-trained) primary teachers entering the system		Probably should be monitored at the start of each school term, or at least at the start of each school year
Indicator 1.2	Teacher absenteeism rate (or achieved annual instruction hours)		To be calculated each term. Reasons for absenteeism would be valuable additional information
Indicator 1.3	Pupil : Textbook/instructional material ratio		
Output 2	Comprehensive provision of classrooms and equipment in rural and urban areas		
Indicator 2.1	Number of classrooms built or rehabilitated and equipped to meet standards		Requires standards to be promulgated and classrooms regularly assessed
Indicator 2.2	Percentage and number of primary schools with inadequate facilities (particularly for health and hygiene, and for girls) in rural and urban areas		Probably should be monitored annually at the start of the school year, when decisions about enrolment are taken by parents
Indicator 2.3	Number of boys and girls being taught in primary schools with inadequate facilities		
Output 3	Effective use of improved standards and policies for gender-sensitive basic education		
Indicator 3.1	Literacy rates among 15-24 year olds, by sex		This will take a little time to improve but is a measure of quality primary education and linked to livelihoods
Indicator 3.2	Performance in maths and science in the National Assessment Test, by sex		
Indicator 3.3	Number of qualified primary school teachers, by sex		The expectation is that this will increase.
Output 4	Improved management of the education system		
Indicator 4.1	Number of education-related strategies and planning frameworks developed and judged fit for purpose		This requires an independent judgement, perhaps by the senior management team

Indicator 4.2	Review and assessment of roles and responsibilities within the education system		This requires an independent judgement, perhaps by the sector performance monitoring team
Indicator 4.3	Service Level Agreements: percentage of sectoral objectives achieved in time		This requires an independent judgement, perhaps by the sector performance monitoring team

3. Health Sector Reform Programme: Strategic Thrusts with a Logical Framework and a Plan of Action, 2004-2007, Federal Ministry of Health, Abuja

3.1. *Introduction*

The text of the strategy is Reference D. To summarise, the performance of the health system in Nigeria is considered to be poor and not able to deliver the health services which the country needs. Without an effective health service, the development aims of the Government are unlikely to be achieved. The strategy proposes action in seven key areas:

1. Stewardship: Improving the role of the Federal Ministry of Health
2. Management: Strengthening the National Health System and its management
3. Disease: Direct action to reduce the burden of disease
4. Resources: Improving the availability of health resources and the way they are managed
5. Access: Improving access to quality health services
6. Awareness: Improving consumer awareness and community involvement
7. Partnerships: Promoting effective partnership and coordination.

The work on logframes carried out as part of the exercises in Module 4 proposed a single unifying logframe to describe the totality of the HSRP, which focuses on areas 1, 2, 3 and 7 as being necessary to achieve the purpose, and suggests that areas 4, 5 and 6 can be lower level activities which in turn support the others.

The logframe is:

HSRP	
Super-goal	To reduce poverty in Nigeria
Goal	To improve the health status of Nigerians
Purpose	To improve the performance of the Nigerian health system (and deliver effective, efficient, good quality and affordable health services)

Output 1	The stewardship role of Government is optimised
Output 2	Federal planning and management is made fit for purpose
Output 3	Reduced burden of disease
Output 4	Partnerships, collaboration and coordination are made effective

Let us now consider each of the levels in turn. It is helpful to start with the Purpose statement, which encapsulates the true aim of the strategy, and for which we expect a specific set of outcomes. Before doing this we shall have a brief look at likely data sources.

3.2. *General comments on data*

It is recommended practice when developing a logframe to state the indicators which will be used for monitoring and the data sources. Unfortunately logframes too often take the sources for granted – phrases such as "source: government statistics" are common – and do not check whether or not they will really provide the information in time. As the remainder of this section will show, health statistics in Nigeria are patchy. Data from surveys are reliable, but can have gaps of 5 years between data points; administrative data are admitted to be poor quality. The implication is that if monitoring is to be taken seriously, resources need to put into collecting the appropriate information.

Key sources noted in the HSRP logframes are:

- National health surveys
- MICS
- DHS and NDHS
- IDRS
- NHMIS
- World Health Report
- Health in Nigeria Report
- Hospital statistics

and we can examine a number of these on the internet.

3.3. *Health statistics*

Consulting the National Bureau of Statistics website (accessed 26 February 2011), the description they give of the state of Health Statistics is not encouraging:

After years of neglect, the health sector has increasingly become a major concern to the Nigerian Government. Improved fertility, better health and the resultant increased life expectancy on the one hand, and population control on the other, are being pursued as mutually compatible programmes in Nigeria today as in most other developing countries. In this direction, Government must rely on adequate health statistics for planning and monitoring its relevant programmes. Data are also needed to estimate health statistical indicators which are required for making geographical (inter-regional and international) comparisons of the health situation in Nigeria. Health statistics are also required for research purposes. Most of the advances made in medicine could not have been possible without health statistics, especially those generated on longitudinal basis. Before the attainment of independence, the items of data on health were collected by the departments of health of the then Regional Governments. Data were obtained from the few general hospitals, infectious diseases hospitals and public health units. Initially, emphasis was on keeping records of reported cases of, and deaths from, communicable diseases such as smallpox, cholera, malaria, tuberculosis, leprosy and yaws. Records of immunisation were also kept. Administrative records of health resources, manpower and vital statistics were among the earliest items of data kept by the regional departments of health. Over the years, there has been a noticeable change in the range of items of data in respect of which records are kept in the health sector, especially on the statistics of primary health care delivery. The Federal Ministry of Health coordinates data collation activities in this sector. With the launching of the National Integrated Survey of Households (NISH) programme in 1981, the National Bureau of Statistics devoted more attention to data collection on some health-related aspects by devoting some modules of the NISH to health, nutritional and demographic matters. In spite

of all these modest efforts, most of which were supported by international agencies, Nigeria's health statistics are in a very bad shape. They are generally uncoordinated, incomplete, unreliable and untimely. This situation has prompted recent efforts of Government and the UNDP to improve on the management of health statistics in the country.

Data can be requested by e-mail from the NBS. The latest publications available from the website include the Directory of Health Establishments 2007 (published 2011); Women and Men in Nigeria 2009; and Social Statistics 2009.

3.4.   *Social statistics*

Social Statistics 2009 has a number of detailed tables containing health statistics. The source of data is primarily from the FMOH, with information in addition from NAFDAC (the National Agency for Food, Drugs and Administration and Control), MICS and the NBS Socio-Economic Survey which appears to be conducted annually. It will be seen that the tables cover many of the key indicators required by the HSRP, though not always precisely as indicated - for example, the desired indicator may be a ratio or proportion of the population, whereas the statistical tables give a headcount. The word "reported" also suggests that no attempt has been made to estimate missing data points. Most of the data are for Nigeria as a whole, some are given at State level. We have indicated tables where more than half the data points are not available.

Nigeria Social Statistics 2009 - Section 4 – Health

Table 4.1 Health Manpower Situation in Nigeria, 2004-2008 (>50% n/a)

Table 4.1a Selected Health Manpower Situation in the State, 2008 (>50% n/a)

Table 4.2a Projected HIV Population by States 2002-2008

Table 4.2b HIV/AIDS Prevalence Rate in Nigeria, 2001, 2003, 2005 & 2008

Module 7 : Case studies (2) : choosing indicators for effective monitoring

Table 4.3 Health Care Facilities by Ownership and State 2004

Table 4.4 Health Care Facilities Population Ratio by State 2000

Table 4.5 Health Care Facilities Population Ratio by State 2004

Table 4.6 Reported Cases of Notifiable Diseases 2002-2008 (>50% n/a)

Table 4.6a Reported Cases of Selected Notifiable Diseases, 2008 (>50% n/a)

Table 4.7 Reported Cases of Deaths from Notifiable Diseases, 2002-2008 (>80% n/a)

Table 4.7a Reported Deaths of Selected Notifiable Diseases, 2008 (>80% n/a)

Table 4.7a cont. Reported Deaths of Selected Notifiable Diseases, 2008 (>80% n/a)

Table 4.8.1 Total Number of Registered Companies by NAFDAC 2001-2006

Table 4.8.2 Number of Registered Products by NAFDAC 2001-2008

Table 4.8.3 Number of Completed Application Forms Received 2002-2008

Table 4.9 Percentage of Women Currently on Family Planning 2001 - 2005

Table 4.10 Percentage of Pregnant Women Receiving Anti-Tetanus Injection by State 2002-2008

Table 4.11 Percentage of Pregnant Women Registered with Clinics by State, 2001-2008

Table 4.12 Crude Death Rate by State 2001-2005

Table 4.13 Percentage of Women Married before 15 Years Old 2001-2005

Table 4.14a Annual Student Enrolment into Professional Health Programmes, 2002 - 2006

Table 4.14b Graduation into the Professions by Year, 2002-2006

Table 4:15 Nigeria Basic Demographic Indicators

Table 4.16 Health Indicators for Nigeria 2005

Table 4.17 Cases of Water Borne Diseases (Diarrhoea, Watery without Blood), 2003 - 2005

Table 4.18 Cases of Water Borne Diseases (Cholera), 2003 – 2005 (>50% n/a)

Table 4.19 Cases of Water Borne Diseases (Diarrhoea, Watery with Blood), 2003-2005 (2005 n/a)

Table 4.20 Cases of Water Borne Diseases Schistosomiasis 2003-2005 (2004 n/a)

Table 4.21 Cases of Water Borne Diseases (Typhoid Fever), 2003 -2005

Table 4.22a Reported Cases of Tuberculosis in Nigeria 2004 - 2008

Table 4.22b Reported Cases of Leprosy in Nigeria 2000 -2008

Table 4.23 Malaria Cases in Nigeria for the Years 2004 - 2008

Table 4.24 Reported Deaths from Malaria, 2002 - 2007

Table 4.25 Reported Cases of Malaria/Anemie Severe 2003-2007

Table 4.26 Reported Cases of Malaria Pregnant Women 15+ 2003-2007

Table 4.27 Reported Deaths from Malaria/Anemie Severe, 2003- 2007 (>50% n/a)

Table 4.28 Reported Death from Malaria Pregnant Women 15+ Year 2004-2008 (>50% n/a)

Table 4.29 Mortality by Gender Age Specific Death Rates by Year and Gender 2002-2006

Table 4.30 Infant Immunisation by State and Year 2003-2005

Table 4.30 cont. Infant Immunisation by State and Year 2006-2007

Table 4.31 Percentage Distribution of Children (One Year or Less) by Type of Immunisation Received, 2008

Table 4.32 Percentage Distribution of Persons Aged 15 Years and Above by Method of Family Planning Used and Sector, 2008

Table 4.33 Percentage Distribution of Deaths in Last 12 Months by Sex of Decease, 2008

Table 4.34 Percentage Distribution of Registered Stillbirth by Year of Registration, 2002-2007 (>80% n/a)

Table 4.35 Percentage Distribution of Birth in the Last 12 Months by State Sector and Weight (kg), 2008

3.5. *Surveys*

Demographic and Health Surveys (DHS) have been carried out in Nigeria four times. At the time this strategy was written (2004) there were survey data available for 1990 and 1999, and the 2003 data were being analysed [source DHS website]. Since then there has been a survey in 2008, and during 2010 a survey specifically on malaria indicators. DHS, therefore, is not suitable for monitoring progress year by year, but will be invaluable for the final evaluation of the HSRP.

Multiple Indicator Cluster Surveys (MICS) are likewise carried out at intervals by UNICEF, using a standard questionnaire which focuses on many of the MDG indicators which would otherwise be unrecorded. Nigeria has participated in all four rounds, 1995, 2000, 2005 (carried out in 2007) and 2010 (being carried out during 2011). The report on MICS3 is in the public domain.

3.6. *Federal Ministry of Health*

The Federal Ministry of Health website was accessed 26 February 2011. The tab *Resources: Health Statistics* leads to two documents but no data. The tab *Departments: Planning, Research and Statistics* merely leads to a description of that department's functions. There is no link to the HMIS as a system – the only reference leads to an NHMIS Policy Document dated December 2006.

The IDSR (Integrated Disease Surveillance and Response) reports are cited in the logframes as being a source of data for many of the disease indicators, but on the FMOH site there is only a 2005 Policy Document. Later documents [Google search] comment that it is not working properly.

3.7. *WHO*

The WHO website was accessed 26 February 2011. The World Health Report 2000 included for the first time a Health Systems Performance Analysis of 191 WHO member states. Unfortunately it has never been repeated. In it, Nigeria was ranked 187th out of 191.

The overall WHO ranking comes from their own "health system performance analysis", an analysis done only once by the WHO for their 2000 World Health Report. It was deemed too expensive an exercise to repeat. The composite index, by which Nigeria was judged to be close to the bottom of the list, was based on quantitative measures of health, such as life expectancy at various ages; disability adjusted life years; expenditure on public health; child and adult mortality; causes of death; and also qualitative measures including the responsiveness of health services (dignity, confidentiality, quality of amenities, access to support, and other factors). These indicators came from specially conducted surveys as well as detailed analysis of existing databases. Therefore the WHO is not necessarily a source for further information, particularly on qualitative measures

3.8. *Purpose indicators*

The Purpose is to improve the performance of the Nigerian health system. The desired aim is to deliver effective, efficient, good quality and affordable health services. What might indicate whether or not the strategy is working?

A good place to start is the introductory chapter which sets out the problems faced in Nigeria. We can highlight a number of statements:

- Nigeria's overall health system performance was ranked 187th out of 191 by the WHO in 2000
- Disease programmes have had little impact since they are implanted within a weak health system
- Primary health care facilities serve only about 5-10% of the potential load
- Public expenditure on health is less than a third of what is recommended
- Federal, State and Local Government Area roles are unclear;

and so on. This analysis led to the identification of the seven strategic areas listed above.

Ultimately, if the services are effective, efficient, good quality and affordable, we would expect health indicators to improve. But that is a measure of the Goal of the strategy. We need to try to find something which is more directly related to the purpose and which bridges the gap between activity and goal.

Note that logframe 1 on page 43, states that a "Health Systems Performance Assessment Report" will be the source of data to measure the effect of improved stewardship. Unfortunately, no further information is available as to who will produce this report and when. As we saw earlier, it will not come from WHO.

Qualitative measures are necessary if we are to judge the effectiveness of the strategy. One possible Purpose indicator could be a change in attitude of Nigerians towards the public health system. If public health provision is indeed improving, one might expect an increase in the number of people using PHC facilities. This could be an indication that services are becoming better quality and more affordable.

For effectiveness and efficiency, we are really looking for signs that the value of the outputs of the system are increasing more than the cost of the inputs. If the strategy is starting to work we would expect the total resources committed to improving health services – from the Federal, State and Local governments, as well as donors and NGOs – to increase. But this money must reach the health centres, hospitals, and the staff, and, as the introduction says in paragraph 11, "the real cost of health services is not known".

Neither of these two options for indicators is satisfactory. We wish to measure a change in attitude, but will have to make do with a headcount of usage. We wish to measure efficiency, but will have to make do with financial commitment. Even these measures will require setting up systems for data collection. But we argue that as projects and programmes become more complex, and involve much more behavioural change, then we also need to invest in

collecting the right information. In this case, a random sample of usage of primary health care centres; and a better system for tracking financial resources are required.

We therefore propose two indicators for Purpose:

1.   The number (or alternatively the percentage increase in the number) of people using PHC facilities;

2.   Total new financial resources committed to the national health system each year (or each quarter) from all sources.

3.9.   *Indicators for Output 1 - Stewardship*

The aim is to optimise the stewardship role of the Federal government. Chapter 2 of the strategy states that there are three basic tasks in stewardship:

1. Setting an explicit health policy development framework, with clearly defined vision, roles and responsibilities, and performance objectives (deliverables) for short and intermediate periods;
2. Exerting influence and ensuring compliance through regulation; and
3. Generating intelligence, that is establishing a reliable data and information base for informed decision-making, monitoring, evaluation and performance assessment.

Perhaps each of these could be recast in a form which would lead us to an indicator of achievement, as follows:

1.1   Number of health policy development frameworks produced, accepted and in operation [note that logframe 1 on page 45 suggests that there will be a number of strategic plans produced at Federal, State and LGA level];

1.2   Instances of non-compliance with regulation;

1.3   Examples of informed use of data and information base.

While 1.1 should be relatively simple to monitor, 1.2 and 1.3 rely on an organisation or individual being made responsible for compiling the information. For 1.2, we are looking for evidence of non-compliance – people will need to be reminded to submit examples to the central point. For 1.3, it should be sufficient to review the development of the database regularly, and survey the main users for their opinion of its quality and utility.

3.10. *Indicators for Output 2 - Management*

The interest of Government (chapter 3) is "to have a health system whose performance is optimal in terms of its goals and functions. The goals of the system are:

1. Provision of good health to the population;
2. Responsiveness to the non-medical expectation of the population, such as respect and confidentiality; and
3. Making financial contributions fair."

These statements are not in a suitable form for turning into indicators as the language is not specific enough. Number 1, in particular, is far too general. [Nor does one provide good health]. Chapter 3 lists a number of strategies for redefining public health functions, but logframe 2 on page 46 only refers (at the highest level) to the same Health Systems Performance Assessment reports as for output 1.

We have suggested for the unifying logframe that the output statement should read:

"Federal planning and management is made fit for purpose"

since a careful reading of the strategy shows that the focus of this component is on structural reform of the FMOH, establishment of a National Hospitals Agency, a new management structure for Federal Health Institutions, the establishment of a National Blood Transfusion System, and the production of a blueprint for reforming the PHC system. The overall aim is to

get strategic planning and management into the right shape to deliver effective health services.

So how can we judge if the national health system has been suitably strengthened and its management performance optimised? Looking at the components of logframe 2 (page 46), we see reference to OOMR (Operations, Organisation and Management Research) Reports as a source of information on whether or not health institutions and their structures are streamlined and functioning correctly. If we assume that these reports are being commissioned, then we could suggest the following:

2.1   Independent assessment of current operational effectiveness of health institutions (the expectation being that year by year a greater number are judged to be operating more effectively);

2.2   Periodic review of defined roles and responsibility of national health structures (the expectation being that gradually these are more clearly defined with less overlap and confusion of roles);

2.3   Assessment of the quality of the blueprint for reforming the PHC (because it is a key part of the HSRP) and its implementation (the expectation being that it will take time to implement, but that regular reports will show positive progress).

3.11. *Indicators for Output 3 – Disease*

We see from chapter 4 (pages 22 and 23) that the burden of disease on the Nigerian population is unacceptably high and getting worse. The main problems lie with primary health care services and emergency obstetric services, but also the public sector response to malaria, TB, HIV/AIDS and other increasingly important risks due to unhealthy lifestyles.

Module 7 : Case studies (2) : choosing indicators for effective monitoring

The strategy involves a mixture of improving PHC services and specific disease control programmes, and on page 23 are a number of suggested indicators by which to monitor this component. We need to find two or three indicators which could represent the whole.

Logframe 3 on page 50 proposes disability-adjusted life expectancy and the prevalence of major diseases (HIV, malaria and TB). DALE, now known as HALE [(Disability-adjusted) Healthy Life Expectancy (at birth)] is a highly technical calculation carried out by WHO specialists and very rarely done in-country. As an indicator it will be useful for measuring long-term change for the whole of Nigeria (assuming a huge improvement in morbidity and mortality statistics, which is by no means certain) but less useful for monitoring progress.

On the other hand, the FMOH and international agencies regularly produce reports on the incidence and prevalence of HIV/AIDS, malaria and TB (through the DOTS programme), and since these are also MDG indicators their quality and timeliness will be better. This, then, is a potential indicator.

Further down page 50, it is suggested that an improvement in PHC services could be measured by "the proportion of health facilities that meet specified national standards". On page 23 this is stated slightly differently, as "the proportion of primary health facilities offering the guaranteed minimum package of services in each state". In addition, it is suggested to monitor "the proportion of the population living within 5km of a health facility. A combination of these statements would make an excellent indicator of availability and quality:

"The proportion of the population living within 5km of a primary health care facility which offers the guaranteed minimum package of services".

For emergency obstetric services (page 51), the other main component of fighting disease, a number of indicators are proposed. Many of these derive from DHS surveys and are less

suitable for monitoring since they are produced on average every five years. We could therefore choose what could be regarded as a "weaker" indicator, such as the second one proposed on page 51 – "number of basic and comprehensive EOC facilities available per 500,000 population". If this could be improved to state that the facilities should meet EOC standards (page 52) we would have a reasonable indicator of availability and quality combined.

Our indicators for Output 3 could therefore be

3.1   Prevalence of malaria, TB and HIV/AIDS;

3.2   The proportion of the population living within 5km of a primary health care facility which offers the guaranteed minimum package of services;

3.3   Number of basic and comprehensive EOC facilities meeting specified national standards, per 500,000 population.

3.12. *Indicators for Output 4 – Partnerships*

The analysis in chapter 8 (pages 31 and 32) refers to the disconnect between public and private health care providers, with the private sector taking an ever greater share but without any coordination of activity. This also applies to the work of international donors, agencies and NGOs.

The aim is to promote effective partnerships, collaboration and coordination through developing an effective policy on public-private partnerships, establishing an effective donor coordination mechanism and promoting "intra-mural practice". [The definition of intra-mural practice is "the delivery of health care services by university faculties or full-time hospital staff conducted within the physical confines of their respective medical centres". It is not clear to a non-medical person precisely what the issue is here, and how this will help the under-

utilisation of facilities in both public and private health facilities, but it is clearly very important.]

What we are therefore trying to monitor is the preparation of policies and their effectiveness. Page 32 suggests monitoring the production of policies, guidelines, legislation and coordination mechanisms plus public-private contracts. This is fairly straightforward though will require someone to keep a record. But more important will be to decide whether these are effective and productive, which is a mainly subjective assessment. Logframe 7 on pages 68 and 69 similarly refers to the existence and availability of documents and mechanisms, the number of agreements signed, but does not try to assess effectiveness.

How might we judge effectiveness? A review of new policies and strategies might look for areas of overlap and also areas which have been overlooked. One would expect the number in both categories to reduce over time. A review of donor coordination might look for concrete results from meetings; for examples of joint working; and for a reduction in the number of new initiatives where donors, or NGOs, or the private sector, are working in isolation. To summarise, we could hope to see a reduction in inefficient working practice.

Based on the above, we can suggest the following indicators:

4.1   Number of good quality policies, guidelines, strategies and coordination mechanisms prepared in time to be useful and which are accepted by major stakeholders;

4.2   Instances of inefficient coordination between donors and government, or public and private sector;

4.3   Proportion of intra-mural arrangements completed compared with numbers of co-existing public-private health facilities.

Indicator 4.3 is taken directly from logframe 7, page 68.

3.13. *Goal indicators*

The goal of the HSRP is "to improve the health status of Nigerians". How this is defined and therefore to be measured is not specified in the strategy. But we can assume that the goal is shared by the Federal Ministry of Health, whose Vision is:

- "to reduce morbidity and mortality rates due to communicable diseases to the barest minimum;
- to reverse the increasing prevalence of non-communicable diseases;
- to meet the global targets on the elimination and eradication of diseases;
- to significantly increase the life expectancy and quality of life of Nigerians."

Let us examine these in turn. Morbidity and mortality rates are slow and complex to calculate and depend on high quality statistical records of disease and death. In countries where the statistical system is weak, with partial records at best, they are often estimated by the WHO. It is unlikely that estimates would be available in time to be useful for monitoring, but they will be crucial for the evaluation and for statements of medium-term change. Note that it is not overall mortality and morbidity rates, but specifically those due to communicable diseases.

Calculating the prevalence of non-communicable diseases (such as cancer, diabetes, and heart disease) also requires good quality statistics. We have seen earlier that there are some data on reported cases, which is a beginning, but for prevalence we need accurate population statistics. Possibly we could monitor reported cases, but as reporting becomes more widespread that could lead to a rise in numbers, even when prevalence may in fact be falling.

Global targets on elimination and eradication of disease refer probably to the MDG targets, which cover HIV/AIDS, malaria and TB, and are subject to the same comments as above – indeed they cover the same topics. We also have some disease indicators for Output 3.

Life expectancy is again often calculated by the WHO. It requires statistics on deaths by age and sex, plus good quality population estimates. It is not normally calculated annually, often being updated only after a census of population.

Quality of life is entirely subjective. A number of organisations compile an index for making annual comparisons between countries or cities, but without additional information there is little we can use here for monitoring.

The logframes in the strategy offer one statement on page 45 which might be of use: "attainment of national/MDG goals in health, including equity, fairness and responsiveness of the health systems".

The MDG health indicators are currently 19 in number, but targets specific to Nigeria are not readily available.

Goal 4: Reduce child mortality	
Target 4A: Reduce by two-thirds, between 1990 and 2015, the under-five mortality rate	4.1 Under-five mortality rate
	4.2 Infant mortality rate
	4.3 Proportion of 1 year old children immunised against measles
Goal 5: Improve maternal health	
Target 5A: Reduce by three quarters, between 1990 and 2015, the maternal mortality ratio	5.1 Maternal mortality ratio
	5.2 Proportion of births attended by skilled health personnel
Target 5B: Achieve, by 2015, universal access to reproductive health	5.3 Contraceptive prevalence rate
	5.4 Adolescent birth rate
	5.5 Antenatal care coverage (at least one visit and at least four visits)
	5.6 Unmet need for family planning
Goal 6: Combat HIV/AIDS, malaria and other diseases	
Target 6A: Have halted by 2015 and begun to reverse the spread of HIV/AIDS	6.1 HIV prevalence among population aged 15-24 years
	6.2 Condom use at last high-risk sex
	6.3 Proportion of population aged 15-24 years with comprehensive knowledge of HIV/AIDS

	6.4 Ratio of school attendance of orphans to school attendance of non-orphans aged 10-14 years
Target 6B: Achieve, by 2010, universal access to treatment for HIV/AIDS for all those who need it	6.5 Proportion of population with advanced HIV infection with access to antiretroviral drugs
Target 6C: have halted by 2015 and begun to reverse the incidence of malaria and other major diseases	6.6 Incidence and death rates associated with malaria
	6.7 Proportion of children under 5 sleeping under insecticide-treated bednets
	6.8 Proportion of children under 5 with fever who are treated with appropriate anti-malarial drugs
	6.9 Incidence, prevalence and death rates associated with tuberculosis
	6.10 Proportion of tuberculosis cases detected and cured under directly observed treatment short courses

NEEDS, the overall development strategy, has four targets for the health sector, from 2003 to 2007

- HIV/AIDS prevalence rate drops from 6.1% to 5%
- Immunisation coverage rises from 39% to 60%
- Access to safe water rises from 64.1% to 70%
- Access to adequate sanitation rises from 53% to 65%.

Targets for infant mortality and maternal mortality were "being developed" (2004).

So – which two do we select?

There is a good argument for selecting the two indicators directly related to health from NEEDS. HIV/AIDS prevalence and immunisation coverage are both highly correlated with the health of the average Nigerian, as the introduction to the HSRP states on page 12 (paragraph 14). For immunisation rates to rise, a lot of other things need to be working more effectively – for example the supply of good quality drugs, accessibility to and utilisation of primary health services, and communication strategies. For HIV prevalence to drop, we can assume that messages about changes in lifestyle are being received positively. We should probably follow the MDG definitions and measure HIV prevalence amongst 15-24 year olds – the population at highest risk – and immunisation against measles as a proportion of 1 year old children.

3.14. *Super-goal indicators*

The HSRP is the health sector's contribution towards NEEDS, the overall Nigerian development strategy. The super-goal is the therefore the reduction of poverty in Nigeria. At this level of the logframe, it is usual to use MDG indicators, since monthly, quarterly or even annual monitoring is not normally required. Clearly the proportion of the population below the national poverty line would be the most appropriate, or below $1 a day at purchasing power parity if national estimates are not available. For a second indicator we could opt for another health indicator, such as infant mortality (which is affected by a number of circumstances – medical, environmental, educational, financial) or an economic one such as employment.

At this level of the logframe (as indeed should be the case at all levels) a round table discussion with stakeholders is the best way to choose an appropriate indicator acceptable to the majority.

3.15. *Summary of unifying logframe and indicators*

Our monitoring framework now looks like this.

HSRP	Narrative	Indicator Source	Comments
Super-goal	To reduce poverty in Nigeria		
Indicator 1	Proportion of population below national poverty line (or $1 a day at 1985 ppp prices)	National Bureau of Statistics for national line, or UNStats website for $1 a day	MDG indicator
Indicator 1	Infant mortality rate	NBS and WHO	MDG indicator
Goal	To improve the health status of Nigerians		
Indicator 1	HIV/AIDS prevalence rate amongst 15-24 year olds	FMOH and DHS	MDG indicator
Indicator 2	Immunisation against measles as a proportion of 1 year old children	FMOH and DHS	MDG indicator

Module 7 : Case studies (2) : choosing indicators for effective monitoring

Purpose	To improve the performance of the Nigerian health system		
Indicator 1	The number (or alternatively the percentage increase in the number) of people using PHC facilities	PHC centre records compiled by FMOH	Will require improved data collection and compilation
Indicator 2	Total new financial resources committed to the national health system each year (or each quarter) from all sources.	FMOH, Federal Ministry of Finance, donors, NGOs	Will require a system for compilation. Could also measure expenditure.
Output 1	The stewardship role of Government is optimised		
Indicator 1.1	Number of health policy development frameworks produced, accepted and in operation	FMOH with State Health Management Boards	
Indicator 1.2	Instances of non-compliance with regulation	FMOH central point	Will require a system for compilation
Indicator 1.3	Examples of informed use of data and information base	FMOH, key users, HSR Implementation Steering Committee	Will require a periodic survey of user opinions
Output 2	Federal planning and management is made fit for purpose		
Indicator 2.1	Independent assessment of current operational effectiveness of health institutions	OOMR reports	The expectation being that year by year a greater number are judged to be operating more effectively
Indicator 2.2	Periodic review of defined roles and responsibility of national health structures	OOMR reports	The expectation being that gradually these are more clearly defined with less overlap and confusion of roles
Indicator 2.3	Assessment of the quality of the blueprint for reforming the PHC(because it is a key part of the HSRP) and its implementation	Independent health expert plus implementation progress reports	The expectation being that it will take time to implement, but that regular reports will show positive progress
Output 3	Reduced burden of disease		
Indicator 3.1	Prevalence of malaria, TB and HIV/AIDS	NBS and DHS	MDG indicators
Indicator 3.2	The proportion of the population living within 5km of a primary health care facility which offers the guaranteed minimum package of services	FMOH, household surveys	Will require a system for compilation
Indicator 3.3	Number of basic and comprehensive EOC facilities meeting specified national standards, per 500,000 population	FMOH, survey of health facilities	Will require a system for compilation
Output 4	Partnerships, collaboration and coordination are made effective		
Indicator 4.1	Number of good quality policies, guidelines, strategies and coordination mechanisms prepared in time to be useful and which are accepted by major stakeholders	HSR Implementation Steering Committee, donors, NGOs	Will require a periodic survey of stakeholder opinions

Indicator 4.2	Instances of inefficient coordination between donors and government, or public and private sector	HSR Implementation Steering Committee, donors, NGOs	Will require a periodic survey of stakeholder opinions
Indicator 4.3	Proportion of intra-mural arrangements completed, compared with numbers of co-existing public-private health facilities.	FMOH MOUs signed	

3.16. *Logframes at lower levels*

It will be recalled from Module 4 that the Nigeria HSRP gave rise to a number of logframes.

The four strategic areas for action, that is to say the Outputs of the unifying logframe, each

have a logframe which links into the unifying logframe, as follows:

	HSRP	Stewardship	Planning	Disease	Partnerships
Super-goal	To reduce poverty in Nigeria	To improve the health status of Nigerians	To improve the health status of Nigerians	To improve the health status of Nigerians	To improve the health status of Nigerians
Goal	To improve the health status of Nigerians	To improve the performance of the Nigerian health system	To improve the performance of the Nigerian health system	To improve the performance of the Nigerian health system	To improve the performance of the Nigerian health system
Purpose	To improve the performance of the Nigerian health system	To optimise the stewardship role of Government	To make Federal planning and management fit for purpose	To reduce the burden of disease	To make partnerships, collaboration and coordination more effective
Output 1	The stewardship role of Government is optimised	Policies	Reform	PHC services	Policy
Output 2	Federal planning and management is made fit for purpose	Legislation	Coordination	EOC services	Efficiency
Output 3	Reduced burden of disease	Communication	Management	HIV+ response	Coordination
Output 4	Partnerships, collaboration and coordination are made effective	Strategies	PHC strategies	Lifestyles	-

We can now consider a framework of indicators for these lower level logframes. We shall choose *Disease* as a working example. *Stewardship, Planning* and *Partnerships* can be done as a practical exercise if desired.

3.17. *Reduced Burden of Disease – an indicator set*

Because we have a hierarchy of logframes, we can use the same indicators where there is overlap. The super-goal statement for *Reduced Burden of Disease* is the Goal statement for the overall HSRP, but can have the same indicators. We can deal similarly with the Goal/Purpose and Purpose/Output 3 indicators. We then immediately have the following:

Disease	Narrative	Indicator Source	Comments
Super-goal	To improve the health status of Nigerians		
Indicator 1	HIV/AIDS prevalence rate amongst 15-24 year olds	FMOH and DHS	MDG indicator
Indicator 1	Immunisation against measles as a proportion of 1 year old children	FMOH and DHS	MDG indicator
Goal	To improve the performance of the Nigerian health system		
Indicator 1	The number (or alternatively the percentage increase in the number) of people using PHC facilities	PHC centre records compiled by FMOH	Will require improved data collection and compilation
Indicator 2	Total new financial resources committed to the national health system each year (or each quarter) from all sources.	FMOH, Federal Ministry of Finance, donors, NGOs	Will require a system for compilation. Could also measure expenditure.
Purpose	Reduced burden of disease		
Indicator 1	The proportion of the population living within 5km of a primary health care facility which offers the guaranteed minimum package of services	FMOH, household surveys	Will require a system for compilation
Indicator 2	Number of basic and comprehensive EOC facilities meeting specified national standards, per 500,000 population	FMOH, survey of health facilities	Will require a system for compilation
Output 1	Improved PHC service delivery		
Indicator 1.1			
Indicator 1.2			

Indicator 1.3			
Output 2	Increased use of modern health facilities for essential obstetric care (EOC)		
Indicator 2.1			
Indicator 2.2			
Indicator 2.3			
Output 3	Effective health sector response to HIV/AIDS and other major diseases established		
Indicator 3.1	Prevalence of malaria, TB and HIV/AIDS	NBS and DHS	MDG indicators
Indicator 3.2			
Indicator 3.3			
Output 4	Healthy lifestyles promoted		
Indicator 4.1			
Indicator 4.2			
Indicator 4.3			

Note that the Purpose statement uses two of the three indicators for Output 3 of the unifying logframe, while prevalence of malaria, etc, is a good measure for Output 3 of this logframe.

We shall not go into as much detail for this logframe. The relevant pages of the HSRP are 50 to 55, where a large number of indicators are suggested. Selecting from these, we could envisage a monitoring framework for *Disease* as follows:

Disease	Narrative	Indicator Source	Comments
Super-goal	To improve the health status of Nigerians		
Indicator 1	HIV/AIDS prevalence rate amongst 15-24 year olds	FMOH and DHS	MDG indicator
Indicator 1	Immunisation against measles as a proportion of 1 year old children	FMOH and DHS	MDG indicator
Goal	To improve the performance of the Nigerian health system		
Indicator 1	The number (or alternatively the percentage increase in the number) of people using PHC facilities	PHC centre records compiled by FMOH	Will require improved data collection and compilation
Indicator 2	Total new financial resources committed to the national health system each year (or each quarter) from all sources.	FMOH, Federal Ministry of Finance, donors, NGOs	Will require a system for compilation. Could also measure expenditure.
Purpose	Reduced burden of disease		
Indicator 1	The proportion of the	FMOH, household surveys	Will require a system for

	population living within 5km of a primary health care facility which offers the guaranteed minimum package of services		compilation
Indicator 2	Number of basic and comprehensive EOC facilities meeting specified national standards, per 500,000 population	FMOH, survey of health facilities	Will require a system for compilation
Output 1	Improved PHC service delivery		
Indicator 1.1	Percentage of PHCs with essential drugs of the right quality always available	Survey of health facilities, regular reporting to health boards	This needs to be monitored in a number of ways, to see whether drug availability is sustained and that the drugs are good quality and not out of date
Indicator 1.2	Percentage of PHCs offering continuous access to the minimum package of services	Survey of health facilities, regular reporting to health boards	Again, the minimum package of services must be sustained once it is established, and the PHC staffed to deliver the services
Indicator 1.3	Percentage of PHC clients satisfied with the quality of the services	Household survey, PHC client service questionnaires	
Output 2	Increased use of modern health facilities for essential obstetric care (EOC)		
Indicator 2.1	Proportion of births attended by skilled health personnel	Hospital statistics and DHS	MDG indicator
Indicator 2.2	Proportion of births which take place in EOC facilities	Hospital statistics, NBS	An alternative indicator could be the MDG one for antenatal care
Indicator 2.3	Neonatal morbidity and mortality	Hospital statistics and DHS	
Output 3	Effective health sector response to HIV/AIDS and other major diseases established		
Indicator 3.1	Prevalence of malaria, TB and HIV/AIDS	NBS and DHS	MDG indicators
Indicator 3.2	Death rates from TB and malaria	FMOH and DHS	MDG indicators
Indicator 3.3	Percentage of people with TB on DOTS	DHS, PHC data, DOTS programme	
Output 4	Healthy lifestyles promoted		
Indicator 4.1	Proportion of population being treated for alcohol-related illnesses	PHC and hospital statistics	Will require improved data collection and compilation
Indicator 4.2	Percentage of the population adopting HIV/AIDS prevention strategies	DHS	DHS data will need to be supplemented by more frequent local surveys. This indicator is also relevant for Output 3
Indicator 4.3	Proportion of the population adopting health-promoting practices	Health surveys	Will need to be specifically organised

This is of course only one possible set of monitoring indicators, and has its faults. Too much depends on systems for data collection and compilation which are not at the moment operational, and as we said earlier, DHS data need to be supplemented by other data to be useful for monitoring.

But – and this is crucial – investment in data collection for monitoring and evaluation is necessary if we want the task to be done properly.

4. Other exercises

For those who would like more practice at developing indicator frameworks for logframes, the following exercises can be attempted. First, modifications to the goal, purpose and output statements should be made to bring them into line with international recommendations. Then suggestions should be made for suitable indicators. Some commentary on the logframes can be found in Module 3.

Exercise 1

Goal	Purpose	Outputs
The people in [state] have improved health and nutrition status	Increased use of quality health, nutrition and sanitation services by the poor	1. Improved access to priority health, nutrition and water and sanitation services in underserved areas 2. Public health management systems strengthened 3. Positive health, nutrition and hygiene practices and health-seeking behaviour of communities improved 4. Improved use of evidence in planning and delivery of equitable health, nutrition and water and sanitation services

Exercise 2

Goal	Purpose	Outputs
To improve the quality of education in [country]	To increase nation wide the access to basic education for children in [country] and the capacity of Ministry of Education to effectively develop Education plans and budgets	1. 2,801,619 primary school going children have access to learning materials and textbooks at 1 to 1 ratio 2. 5300 SDCs (school development committees) revitalised and functional leading to increased parents participation in school management. 3. National education recovery plan developed including a prioritisation and costing plan

Exercise 3

Goal	Purpose	Outputs
To improve health &	To meet emergency and	1. Improved sustainable access to safe, reliable water

well-being of vulnerable households in [state]	recovery needs of conflict affected communities in [state]	supply and to safe sanitation facilities
		2. Improved health and hygiene related awareness and behaviours at community level
		3. Improved access to nutritional support
		4. Increased access to education infrastructure
		5. Improved food security and tree coverage

References

A   Republic of The Gambia Education Sector Medium Term Plan 2008-2011

D   Health Sector Reform Programme: Strategic Thrusts with a Logical Framework and a

Plan of Action, 2004 – 2007, Federal Ministry of Health, Abuja

The following websites have been consulted for this module.

www.uis.unesco.org	UNESCO Institute for Statistics, glossary of education terms
www.measuredhs.com	DHS surveys
www.data.un.org	UN database
www.unicef.org	MICS surveys
www.who.int	WHO, World Health Report, glossary of health terms
www.gbos.gm	Bureau of Statistics, The Gambia
www.edugambia.gm	Department of State for Basic and Secondary Education, The Gambia
www.nigerianstat.gov.ng	National Bureau of Statistics, Nigeria
www.fmh.gov.ng	Federal Ministry of Health, Nigeria

Better Monitoring for Better Evaluation

Module 8: – Setting-up and Using a Monitoring System:

How to improve decision making

Contents

1. Setting up a management system

1.1. *Why Plan for Monitoring and Evaluation?*

Effective and timely decision making requires information from regular and planned monitoring and evaluation activities. Planning for monitoring and evaluation must start at the time of programme or project design, and they must be planned together. While monitoring provides real-time information on ongoing programme or project implementation required by management, evaluation provides more in-depth assessments. The monitoring process can generate questions to be answered by evaluation. Also, evaluation draws heavily on data generated through monitoring, including baseline data, information on the programme or project implementation process, and measurements of progress towards the planned results through indicators. Planning for monitoring must be done with evaluation in mind: The availability of a clearly defined results or outcome model and monitoring data, among other things, determine the 'evaluability'[7] of the subject to be evaluated.

---

[7] Evaluability can be defined by clarity in the intent of the subject to be evaluated, sufficient measurable indicators, accessible reliable information sources, and no major factor hindering an impartial evaluation process.

1.2.   *Monitoring and Evaluation Framework*

A clear framework, agreed among the key stakeholders at the end of the planning stage, is essential in order to carry out monitoring and evaluation systematically. This framework serves as a plan for monitoring and evaluation, and should clarify:

What is to be monitored and evaluated

The activities needed to monitor and evaluate

Who is responsible for monitoring and evaluation activities

When monitoring and evaluation activities are planned (timing)

How monitoring and evaluation are carried out (methods)

What resources are required and where they are committed

In addition, relevant risks and assumptions in carrying out planned monitoring and evaluation activities should be seriously considered, anticipated and included in the M&E framework.

In general, the M&E framework has three main components:

1.2.1.   Narrative component

This describes how the partners will undertake monitoring and evaluation and the accountabilities assigned to different individuals and agencies. For example, at the UNDAF or national result level, it is necessary to engage with national monitoring committees or outcome level groups (e.g. sector arrangements) as well as with UN interagency monitoring working groups. If these do not exist, there might be a need to establish such structures for effective monitoring and evaluation. In addition the narrative should also reflect:

a. Plans that may be in place to strengthen national or sub-national monitoring and evaluation capacities

b. Existing monitoring and evaluation capacities and an estimate of the human, financial and material resource requirements for its implementation

1.2.2.   Results framework

This should be prepared in the planning stage as described below

1.2.3. Planning matrices for monitoring and evaluation

These are strategic and consolidate the information required for monitoring and evaluation for easy reference. The planning matrix for monitoring in the following Table is illustrative for UNDP and could be used at the country, regional and global programme level to determine what needs to be monitored. (A completed example of this Table is given in the subsequent chapter) This matrix should be adapted as determined by local circumstances and conditions. In some cases, the columns could be modified to cover results elements such as outcomes, outputs, indicators, baselines, risks and assumptions separately. The need for an M&E framework applies for both programmes and projects within a programme. Therefore both programmes and projects should develop M&E frameworks in their planning stages. The project-level M&E framework should cascade from the programme level M&E framework and could contain more detailed information on monitoring and evaluation tasks that apply specifically to respective projects. Conversely, the programme-level framework builds upon the project-level frameworks. Monitoring and evaluation activities should be seen as an integral component of programme and project management. They take place throughout the programme and project cycles and should be reviewed and updated regularly (at least annually, for example at the time of annual reviews).

1.3. *Planning matrix for monitoring*

Planning matrix for monitoring[24]							
Expected Results (Outcomes & Outputs)	Indicators (with Baselines & Indicative Targets) and Other Key Areas to Monitor	M&E Event with Data Collection Methods	Time or Schedule and Frequency	Responsibilities	Means of Verification: Data Source and Type	Resources	Risks
Obtained from Development plan and results Framework.	From results framework. Indicators should also capture key priorities such as capacity development and gender. In addition. other key areas need to be monitored, such as the risks identified in the planning stage as well as other key management needs.	How is data to be obtained? Example: through a survey. a review or stakeholder meeting, etc.	Level of detail that can be included would depend on the practical needs.	Who is responsible for organizing the data collection and verifying data quality and source?	Systematic source and location where you would find the identified and necessary data such as a national Institute or DevInfo[8].	Estimate of resources required and committed for carrying out planned Monitorin activitie activities	What are the risks and assumptions for carrying out the planned monitoring activities? How may these affect the planned monitoring events and quality of data?

The content of this table should become more evident in the detailed example on next page

---

[8] DevInfo does not seem to be an approproiate source of verification but a very useful source for presenting, feeding them into the decision making

process and harmonizing the data

1.4. *Example of a planning matrix for monitoring*

Illustrative example of planning matrix for monitoring: Enhanced capacity of electoral management authority							
Expected Results (Outcomes & Outputs)	Indicators (with Baselines & Indicative Targets) and Other Key Areas to Monitor	M&E with Data Collection Methods	Time or Schedule and Frequency	Responsibilities	Means of Verification: Data Source and Type	Resources	Risks
Outcome 1: Enhanced capacity of electoral management authority to administer free and fair elections 1.1. Advocacy campaign aimed at building consensus on need for electoral law and system reform implemented 1.2. Electoral management authority has adequate staff and systems to administer free and fair elections 1.3.Training programme on use of new electoral management technology designed and implemented for staff of electoral management authority	Public perception of capacity of electoral management authority to administer free and fair elections (disaggregated by gender, population group, etc.) **Baseline:** 40% of public had confidence in electoral management authority as of 2008 (50% men, 30% women, 20% indigenous populations) **Target:** 70% of overall population have confidence in electoral management authority by 2016 (75% men, 65% women, 60% indigenous populations)	1. Surveys 2. Annual Progress Reviews 3. Joint field visits to five regions	1. LAN surveys will be completed six months prior to the completion of activities 2. Progress reviews on achievement of all connected outputs will be held jointly in the fourth quarter 3.Two field visits will be held prior to	1. National Office of Statistics will commission survey; external partners, UNDP, and the World Bank will provide technical resources as needed through their assistance for capacity development 2. Progress Reviews will be organized by Elections Authority 3. Field visits will be organized by Elections Authority; Elections Authority will ensure meetings with a representative	1.1 Data and analysis of surveys will be available in (a) report for public And (b) on websites of National Office of Statistics and Elections Authority 2.1 Annual Progress Reports 2.2 Minutes	1. Resources estimated at USD 0.2 million for the survey will be provided by the European Union 2. Resources for M&E activities will be made available in World Bank assistance project 3. Cost of external partners' participation will be met by each	1.It is assumed that capacity development activities within National Office of Statistics required for carrying out the survey will be completed one year in advance to actual survey; if there are

Outcome 2: Increased participation by women and indigenous populations in national and local electoral processes in five regions by 2016 2.1. Revised draft legislation on rights of women and indigenous populations to participate in elections prepared	Percentage of eligible women registered to vote in five regions **Baseline:** 30% of eligible women registered in the five regions as of 2008 **Target:** 60% registration of eligible women in the five regions by 2016	the final survey and three more afterward th continued	cross section of stakeholders; at least two external partners will participate in a given joint field visit	of Annual Progress Reviews 3. Records of joint field visits will be available on website of Elections Authority	respective partner. Other logistical costs will be funded from World Bank project tbc	delays, then a private company could be contracted to carry out the survey tbc

1.5. *Exercise: Design a Planning matrix for monitoring*

Finally us repeat some parameters essential for building a monitoring system

1.6.   *Resources for Monitoring and Evaluation*

Inadequate resources lead to poor quality monitoring and evaluation. To ensure effective and quality monitoring and evaluation, it is critical to set aside adequate financial and human resources at the planning stage. The required financial and human resources for monitoring and evaluation should be considered within the overall costs of delivering the agreed results and not as additional costs.

Financial resources for monitoring and evaluation should be estimated realistically at the time of planning for monitoring and evaluation. While it is critical to plan for monitoring and evaluation together, resources for each function should be separate. In practice, each project should have two separate budget lines for its monitoring and evaluation agreed in advance with partners. This will help UNDP and its partners be more realistic in budgeting. It will also reduce the risk of running out of resources for evaluation, which often takes place towards the end of implementation. Monitoring and evaluation costs associated with projects can be identified relatively easily and be charged directly to the respective project budgets with prior agreement among partners through inclusion in the project budget or Annual Work Plan (AWP) signed by partners. Sourcing and securing financial resources for monitoring and evaluation of outcomes or programmes can pose additional challenges, as there is not one project where these costs can be directly charged. The most commonly observed financing mechanism is to draw resources together from relevant projects. Some additional possibilities include:

Create a separate monitoring and evaluation fund, facility or project associated with an outcome or a programme to which all the constituent projects would contribute through

transfer of some project funds. This facility could be located in the same entity that manages the outcome or programme.

Mobilize funds from partners directly for an outcome or programme monitoring and evaluation facility.

Allocate required funds annually for each outcome on the basis of planned costs of monitoring and evaluation from overall programme budget to the facility or fund.

It is important that partners consider the resources needed for monitoring and evaluation and agree on a practical arrangement to finance the associated activities. Such arrangements should be documented at the beginning of the programme to enable partners to transfer necessary funds in accordance with their procedures, which could take considerable time and effort

Human resources are critical for effective monitoring and evaluation, even after securing adequate financial resources. For high-quality monitoring and evaluation,
there should be:

Dedicated staff time: For effective monitoring and evaluation, staff should be dedicated for the function. The practices of deployment of personnel for monitoring vary among organizations.

Skilled personnel: Staff entrusted with monitoring should have required technical expertise in the area.

Each monitoring and evaluation entity that functions at different levels, for example at the project, programme or outcome level, should have a clear ToR outlining its role and responsibilities. In general, these responsibilities should include:

Setting up systematic monitoring frameworks and developing an evaluation plan

Meeting regularly with key partners and stakeholders to assess progress towards achieving the results

Conducting joint field monitoring and evaluation missions to assess achievements and constraints

Identifying any lessons or good practices

Reflecting on how well the results being achieved are addressing gender, and the interests and rights of marginalized and vulnerable groups in the society

Identifying additional capacity development needs among stakeholders and partners

Reporting regularly to the lead individuals or agencies for the particular result areas and seeking opportunities to influence policy and decision-making processes

1.7. *Engagement of Stakeholders In Monitoring And Evaluation*

The engagement of stakeholders enlisted during planning and described before continues to be relevant for monitoring and evaluation stages for the following reasons:

The stakeholders, who set the vision and the prioritized results to realize that vision during the planning stage, have the best ideas on how the results would continue to remain relevant to them. They must therefore be involved in identifying the **information or feedback that is needed during implementation**, which determines the parameters for monitoring and evaluation.

Having set the vision, priority results and initial parameters for monitoring and evaluation, the key stakeholders are best placed to ensure that the programmatic initiatives planned would deliver what was intended and the way it was intended.

Stakeholder participation in monitoring and evaluation can produce effective communication for various other objectives. These include: facilitate communication of 'early wins' to

increase support and enlist engagement of those who are not yet engaged, ensure access of early products and services of initiatives for intended beneficiaries, mobilize additional resources to fill resource gaps, and ensure effective use of lessons learned in future decision making.

Stakeholder participation throughout the programming cycle ensures ownership, learning and sustainability of results. **Continued stakeholder participation in monitoring and evaluation cannot be assumed. It must be institutionalized.**

At the higher levels of results (national goals, sector goals and outcomes), key stakeholders should typically form sector-wide or inter-agency groups around each major outcome or sector. Whenever there are existing national structures such as sector-wide coordination mechanisms, the United Nations and UNDP should ideally engage them and participate in these rather than setting up parallel systems. Sectoral or outcome-level coordinating mechanisms should not be a United Nations or UNDP management arrangement, but an existing national structure that is already charged with the coordination of the sector from a development perspective within the national context. These groups should have adequate capacity to be responsible for the following:

Agree on an M&E framework for the outcomes and oversee their implementation. They ensure continuous outcome assessment and can enhance progress towards results. Promote partnerships and coordination within a single shared outcome. All projects that are generating relevant outputs to the corresponding outcome should be included in the outcome group to ensure inclusive discussions. This gives partners a common vision of the outcome to which different projects or outputs are contributing. Ensure synergy and coordination by reinforcing a common strategy among partners working towards common results.

Monitor and evaluate, where appropriate, the achievement of outcomes and their contribution to national development goals. Outcome-level mechanisms are expected to determine who is responsible for monitoring and data collection, how often it will be collected, who will receive it and in what form. The results frameworks and the M&E framework serve as the basis for joint monitoring and evaluation by these groups.

Carry out, participate in, and assure the overall quality of project, outcome, thematic and other types of reviews and evaluations and ensure that the processes and products meet international standards. (see quality assessment procedures)

Ensure effective use and dissemination of monitoring and evaluation information in future planning and decision making for improvements.

Capacities for monitoring and evaluation, like for most technical areas, exist on three levels: the enabling environment, the organizational level, and the individual level. Capacities at these levels are interdependent and influence each other through complex co-dependent relationships. Change in capacity generally occurs across four domains: institutional arrangements, including adequate resources and incentives; leadership; knowledge; and accountability mechanisms.

Addressing only one of these levels or domains in a programme or project is unlikely to result in developing sustainable monitoring and evaluation capacities. Therefore, an outcome group needs to take a more holistic view in identifying and addressing the capacities needed to monitor and evaluate the results being pursued.

Benchmarks for the three levels and four domains mentioned above are limited. However, the subsections below offer possible lines of questioning for the preliminary assessment. The insights generated by these questions and others may help a programme team formulate a capacity development response.

1.8. *Institutional Arrangements*

Is there a documented institutional or sector programme monitoring and evaluation policy that clarifies the mandates of monitoring and evaluation entities and programme or project teams, their responsibilities, and accountability measures for effective data collection and data management of public programmes or projects?

Does the institutional and sector policy mandate require: establishing standard tools and templates, aligning organizational data with the national data collection and management, defining standards for monitoring and evaluating skills, and ensuring proper training?

Are sufficient resources, including availability of skilled staff and financial resources, allocated for monitoring and evaluation activities in respective monitoring and evaluation entities? Do monitoring staff have proper statistical and analytical skills to compile and analyse sample and snapshot data?

Is there an independent evaluation entity? Is the institution responsible for evaluation truly 'independent' from management and subject to evaluation? What is the reporting line of those responsible for carrying out evaluations? What mechanisms are there to safeguard the independence of the evaluation function?

1.8.1. Leadership

Does high-level management support evidence-based decision making throughout the organization?

1.8.2. Knowledge

Can high-quality information be disaggregated by relevant factors (such as gender, age and geography) to assess progress and analyse performance?

Do the respective monitoring and evaluation entities have access to all relevant programme or project information to be gathered? Do the stakeholders have access to data collected and analysed (for example through the Internet)?

Do the monitoring and evaluation entities have easy-to-understand formats for data collection and reporting? Is there a systematic and documented process of ensuring data quality control at all levels of collection, analysis and aggregation?

Is there sufficient evaluation technical expertise in the national system? Are there national professional evaluation associations?

1.8.3. Accountability

Can the information from the monitoring and evaluation entities be provided to decision makers and other relevant stakeholders in a timely manner to enable evidence-based decision making?

Based on the above considerations and the insights generated from a high-level capacity assessment, one of four broad approaches would be selected to meet the monitoring and evaluation requirements of the results being pursued (see next Figure). This high-level capacity assessment may also lead to more in-depth capacity assessments for particular areas. It may be important for the sector-wide or outcome group to document the analysis from the following simplified Figure a simple capacity development matrix (see following Table). This matrix can help determine what monitoring and evaluation facilities exist in national partner institutions that can be used and identify gaps. The last column could be used to indicate how capacity development efforts—including detailed capacity assessments— may be addressed through other programmatic support, when relevant national demand and need arise.
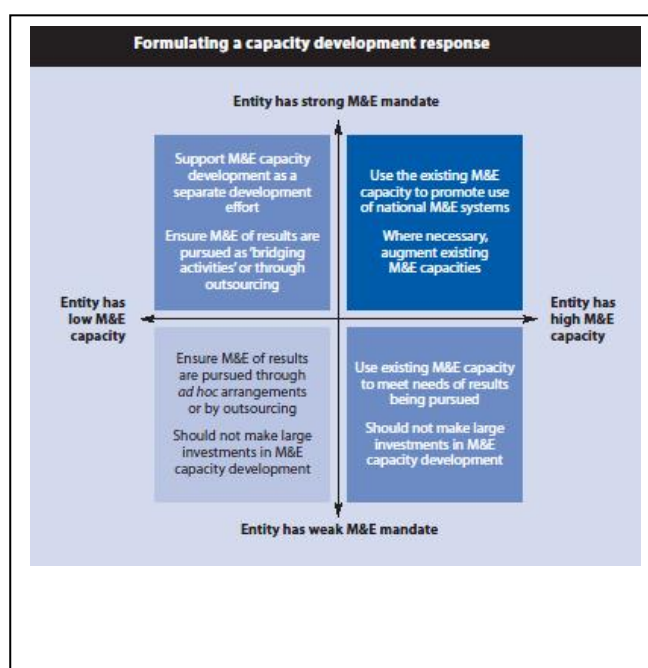
Monitoring and evaluation capacity matrix

Key Partner or Stakeholder of the Outcome Group Contributing to Result	Specific Component of Result or Outcome for Which the Partner is Directly Associated	Existing M&E Mechanisms and Capacities of Partner (institutional arrangements, leadership, knowledge, accountability)	Potential Areas for Developing M&E Capacities of Partner in Line with Its Mandate	Recommended Action for Developing M&E Capacities
Elections Authority	■ Organizing progress reviews, field visits ■ Collection and analysis data ■ Reporting	Limited to Headquarters level only.	Field monitoring, especially skills at the regional level to assess inclusion of disadvantaged and those in remote locations.	Initial capacity development support should be focused on developing monitoring skills pertaining to achieving the outcome. Funds available within the outcome may also be used to carry out a capacity assessment for the Elections Authority.
National Office of Statistics	All surveys will be completed by National Office of Statistics.	National Office of Statistics is a key national institute that is expected to provide high quality national surveys, analyses and reporting of findings.	Capacity development of National Office of Statistics is a national priority.	The Outcome Group should promote a national effort to develop capacity of National Office of Statistics for conducting, analysing and reporting on surveys.
Monitoring and Evaluation Division, Ministry of Planning	Government unit responsible for monitoring and evaluating major development projects and coordination of sector-level monitoring and evaluation (including the election project) at the national outcome level, and to build the national capacity in monitoring and evaluation.*	Monitoring and Evaluation Division is politically independent and is staffed with civil servants competent in monitoring and evaluation.	Monitoring and Evaluation Division has never worked directly with staff members of the Election Authority or National Office of Statistics regarding monitoring and evaluation in this particular area. This is at high risk to be politicized.	Support the efforts of the Monitoring and Evaluation Division to train the Election Authority Electoral Commission staff and National Office of Statistics staff on the development of specific indicators, baselines and targets and data collection methods for the work of the Elections Authority. Support the efforts of the Monitoring and Evaluation Division to promote the culture of evaluation within the Elections Authority.

As these measures are essential for a sustainable management system based on a result based monitoring framework, the continuing effort to keep the system working efficiently cannot be described as a method, it is an ongoing process supported by the related organization and stakeholders.

2. Examples of Reporting the information

2.1. *Why to discuss indicators here?*

Focusing on the most commonly used measures promotes comparability across regions and groups. These include the headcount index, poverty gap, and squared poverty gap. Others are measures of equality/ inequality like Gini (Lorenz curves) or even better Theil-Index. These are all quite well known tools and have been discussed in the indicator section. Here it will be to show some examples in the context of an (attempt of) evaluation. Another reason for this chapter is to point to possible supplement or completion of these established indicators, especially incorporating qualitative indicators into the context of evaluation. Another reason for this chapter is to point to possible supplement or completion of these established indicators, especially incorporating qualitative indicators into the context of evaluation. These examples can be incorporated into the chapter of presentation of performance data later (chapter 3.3)

2.2. *Presenting Poverty Indicators*

Just to recall the most common poverty indicators:

The headcount numbers the poor (in relation to the total population)

The poverty gap shows the relative distance (on average) the poor have to travel in raising incomes to leave poverty behind.

The squared poverty gap shows the distribution of poverty among the poor, the higher it is the bigger are the differences among the poor

The Watts Index measures in log-terms the differences between the poor and the poverty lines

All these values are usually displayed in percentages (%) where the interpretation of a

percent in the case of all but the headcount is not obvious.
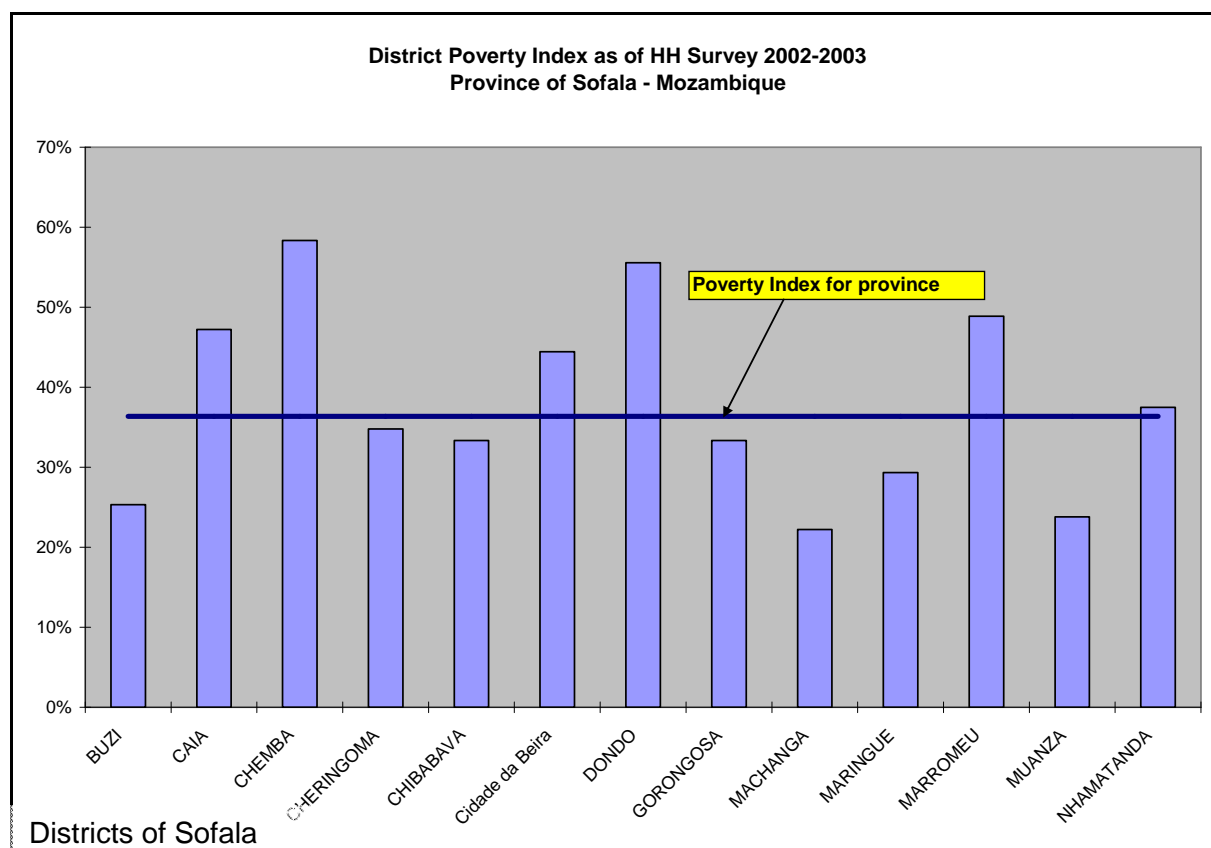
Graph 8-1

A display of provincial poverty headcount (as above) is quite straightforward and increase

and decrease can also be displayed quite easily like placing two bars for each province next

to each other.

A presentation of mapped indices is sometimes preferable because it relates spatial

information to the poverty indicators, pointing to the question: Why is poverty less in the

central provinces and highest in some provinces in the south. A combination of the two might

be a best solution, the interpretation of poverty gap and squared gap and Watt's poses a

serous problem in interpretation to the public. Because of these difficulties these indices are

often left out and their low profile in the MDGs as in Target 1A is often just reduced to the

Module 8: – Setting-up and Using a Monitoring System: How to improve decision making

headcount (population living on 1 / 1.25 US$) a day. This indicator allowing at least some

analysis about inequality or possibilities to overcome them is often tacitly ignored in the
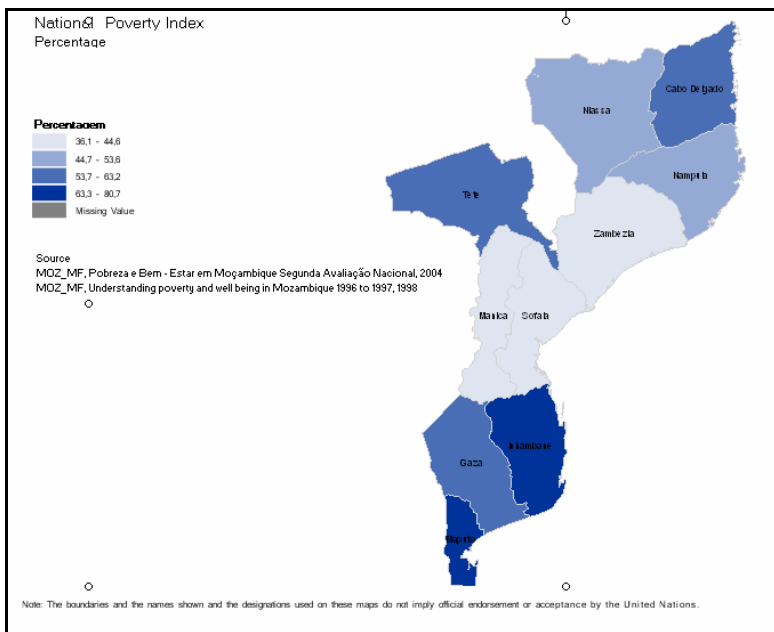
presentation of MDGs. This should be otherwise in national plans if the objective would be to reduce regional discrepancies of poverty and / or economical well-being.

The headcount is easy to grasp even by a rather not analytically minded audience (40% are poor is easier to understand than the gap ratio is 10% below the poverty line).

However for comparison of equality and inequality the next set of indicators are useful

2.3.   *Presenting Inequality Indicators*

The classical presentation of inequality is the Lorenz curve and the Gini coefficient

representing the percentage of the space covered by the Lorenz curve under the diagonal
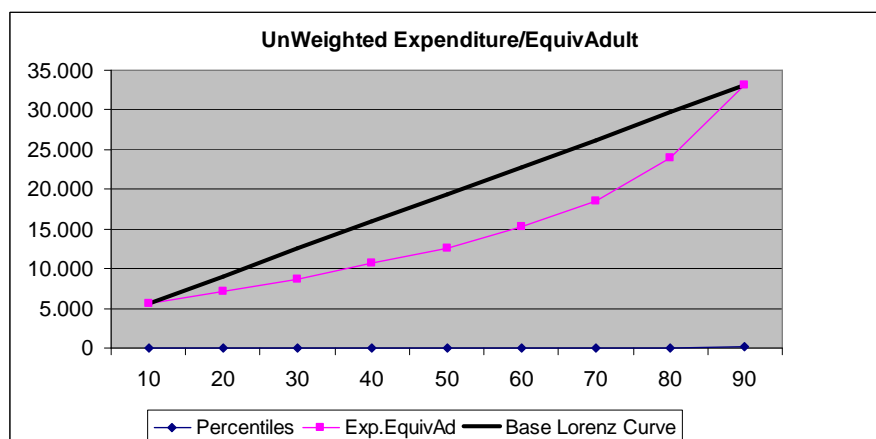
comparing population groups and their expenditures / income / proportion of wealth. The

value differs between 0 (perfect equality) and 1 (one individual owns, spends everything). The numbers are often easier to compare than the proper Lorenz curve, meaning higher number stands for more inequality.

The percentages could likewise be presented as bars or time lines.

Another classical indicator of inequality is the Engel's coefficient. Usually the relation of food expenditure to the total expenditure is called the Engel's coefficient and Engel's Law claims in short, that this coefficient rises the poorer the household is. The coefficient being a percentage allows a presentation as before.

2.4.   *Expected Results*

Results of Monitoring development. The expected results are indicators of development for a year. The value is displayed in percentage varying between -100% (extremely negative and +100% (extremely positive).This will be displayed in form of graphics and tables for the indicated region and time period in this case for 12 districts over 3 years. There is no base line although the data of the first year (2005) could be used as such and relative increase / decrease is not displayed here but could easily been done. What was used is a qualitative questionnaire. The questionnaire is very simple can be answered by the head of household or the substitute and would not required neither extended length of permanence in the household nor extended processing requirements.

It contains 18 questions with 5 possible answers (usually from *deteriorated very much*.. to… *improved very much*). Questions are coded (-2,-1,0,1,2) accordingly and the sum is divided by (2* number of questionnaires). This technique has been applied to other longitudinal surveys and allows monitoring development over time. It permits to get almost instant responses with quick tendency results. The qualitative results are elaborated and indexed allowing immediate information about improvement or deterioration in the questioned area.

The results of these interviews are opinions and not facts. Therefore the qualities of results have to be controlled by quantitative surveys, preferably the Census, HH surveys or CWIQ (Core Welfare Indicator Questionnaire). By using a survey over a longer period, both quantitative and qualitative surveys can complete each other.

Graph 8.2 compares results for subsequent years. Although no proper baseline is given, it is implied and suggested to compare the data of following years to the ones of the first year of the survey. One can assume the zero line for the year preceding the annual surveys (2004) as a baseline because the questions in the survey refer to observed changes during last year. More critical would be to establish a baseline from a separate source like a household survey because both surveys ask different questions and cannot be easily compared.

SMD 2005/2006/2007 Development Indices
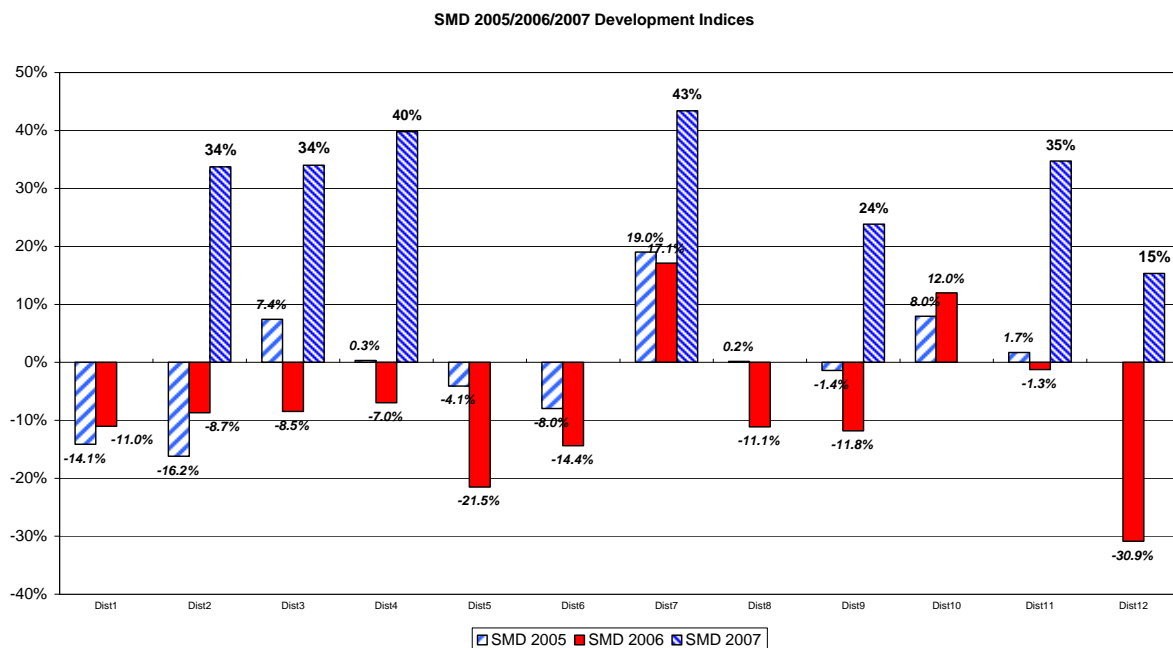
Graph 8-2

2.5.   *Presenting Welfare Indicators*

Revolutionary Measures

Using Tunisia and Egypt as benchmarks, measures for other areas are shaded if they meet or exceed those thresholds. (In the case of median age and level of democracy, lower numbers are shaded.)

Data not available for all categories or all areas	Median age	Unemployment rate — Most recent estimates	Income inequality — (Gini Index) Higher numbers represent more income inequality	Spending on Food — On food consumed at home, as a percent of household spending, 2008	Both from the Economist Intelligence Unit's "Democracy Index 2010"		Internet Penetration — Users as a percentage of the population, 2010
					Level of Democracy (Scale of 1 to 10)	Regime type	
Tunisia	29.7	14.0	40.0	35.8	2.8	Authoritarian	34.0
Egypt	24.0	9.7	34.4	38.3	3.1	Authoritarian	21.2
Other North African							
Algeria	27.1	9.9	35.3	43.8	□.□	Authoritarian	□□.□
Morocco	26.5	9.8		40.3	□.□	Authoritarian	
Libya	24.2	30.0			1.9	Authoritarian	□.□
Sudan	18.4	18.7			2.4	Authoritarian	□□.□
Middle Eastern							
Jordan	21.8	13.4	39.7	40.8	□.□	Authoritarian	27.2
Iran	26.3	14.6	44.5	□□.□	1.9	Authoritarian	43.2
Jemen	17.9	35.0	37.7		2.6	Authoritarian	□.□
Oman	23.9	15.0			2.9	Authoritarian	41.7
Saudi Arabia	24.9	10.8		□□.□	1.8	Authoritarian	38.1
Turkey	28.1	12.4	41.0	□□.□	□.□	Hybrid	45.0
Kuwait	26.4	□.□	□.□	□□.□	□.□	Authoritarian	39.4
Bahrain	□□.□	15.0		□□.□	□.□	Authoritarian	88.0
Lyria	21.5	□.□			2.3	Authoritarian	□□.□
Israel	29.3	□.□	39.2	□□.□	□.□	Flawed democracy	71.6
Katar	□□.□			□□.□	3.1	Authoritarian	51.8
Gaza Strip	17.5	40.0			□.□	Hybrid	
West Bank	20.9	16.5			□.□	Hybrid	□□.□
Irak	20.6	15.3			□.□	Hybrid	□.□
Lebanon	29.4				□.□	Hybrid	24.2
United Arab Emirates	□□.□	□.□		□.□	2.5		75.9
Zyprus	□□.□	□.□	□□.□		□.□	Flawed democracy	39.3
For Comparison							
United States	□□.□	□.□	45.0	□.□	□.□	Full democracy	77.3

Source: Data for median age, unemployment rate and income inequality from the CIA's "The World Factbook" (U.S. unemployment rate from the Bureau of Labor Statistics); food spending data from the U. S. Census Bureau's "The 2011 Statistical Abstract"; level of democracy and regime type data from the Economist Intelligence Unit's "Democracy Index 2010" ; and, Internet penetration data from Internet World, internetworldstats.com

Now combining some of these indicators, for being short we allow all kind of creativity. For being short we just call them welfare indicators to explain that a complex phenomenon like well-being or welfare can hardly be explained by one indicator. An example below is the Revolutionary index (from NYTimes – 02.21.2011). The most obvious way to present this is showing the different indices side by side (like below). The disadvantage of this is that the different individual indices have to be explained separately and no common (weighted) indicator is selected. If however a weighted indicator would be presented than the choice of weights and the "meaning" of the indicator requires lengthy explanations. Examples are the Human Development Index (HDI), the

Another example is a measure of well-being on district level as presented below (from [10])

This Index combines economical well-being with satisfaction of health support, education and the performance of the local administration with qualitative methods.

SMD Indices -	2005	2006	2007
Buzi	-14%	-11%	
Caia	-16%	-9%	34%
Chemba	7%	-8%	34%
Cheringoma	0%	-7%	40%
Chibabava	-4%	-21%	
Dondo	-8%	-14%	
Gorongosa	19%	17%	43%
Machanga	0%	-11%	
Maringue	-1%	-12%	24%
Marromeu	8%	12%	
Muanza	2%	-1%	35%
Nhamatanda		-31%	15%

System of Monitoring Development in the Province of Sofala (Mozambique)

If this can be called "District Development" is debatable but as an indicator of decentralized development it is certainly an early and complex lead of poverty and well-being.

2.6.   *Knowledge Products and Dissemination*

Monitoring and Evaluation has no justification in itself but has to result in relevant, timely, usable and available knowledge products

Knowledge products can take many different forms depending on the audience and their information needs. For meaningful learning and knowledge sharing, knowledge products should be of high quality with a clearly identified audience and purpose. The characteristics of a good knowledge product, including a good publication, are listed in Box 43. Keeping these characteristics in mind before the starting analysis or preparing a knowledge product will help organize the evidence in an orderly fashion.

Practical Steps for Developing Knowledge Products and Dissemination

The **dissemination** is as important as the development of knowledge products. Only an efficient system of dissemination will ensure that the target recipients receive the monitoring and evaluation feedback that is relevant to their specific needs. Some of the most commonly applied dissemination methods for monitoring and evaluation products include: printed reports, HTML or PDF copies of the products shared on the internal and external Internet sites and through e-mail messages and list-serves, and CD-ROMs. The media can be a powerful partner in disseminating findings, recommendations and lessons from evaluation. In many countries, the media has played a critical role in advocating for accountability and addressing sensitive issues. The following are practical steps for developing knowledge products from monitoring and evaluation and disseminating them.

Step 1: Identify target audiences and their information needs

Some of the commonly identified key target audiences for evaluation reports and knowledge products are the following:

Government counterparts who may or may not be directly involved in the project under evaluation but can facilitate the policy changes recommended by the evaluation or otherwise aid in the country-level advocacy

Development partners, other UN organizations, NGOs, and academic and research institutions

Other networks of evaluators (for example, a national evaluation association) Those responsible for knowledge sharing and dissemination should assess the information needs of the various groups, including when the information is most needed and make the information available in a manner that is appropriate for the technical and functional needs of the target audience.

Step 2: Collect stakeholder contact information

The success of every dissemination effort is highly dependent on the recipient contact information gathered during the monitoring and evaluation processes. For example, the

evaluation team members meet with key stakeholders and national counterparts who, regardless of their degree of involvement in the evaluation topic, constitute a critical audience and should be informed about the knowledge generated from evaluation. The contact information of these individuals should be gathered by the evaluation team and shared with those responsible for disseminating and sharing the knowledge.

Step 3: Determine types of products that meet the audience's information needs

In addition to publishing information from regular monitoring reports61 and evaluation reports, a mix of knowledge products can be developed to meet the information demand of different groups. A systematic assessment of the needs and demand for specific products among targeted audiences can be undertaken to ensure the relevance and value of the products.

Step 4: Identify language requirements per product and audience

In order to optimize the impact of knowledge sharing and dissemination efforts, knowledge products should be translated into local languages whenever possible. If resources are limited, the commissioning unit may determine language requirements per knowledge product or per audience group. At a minimum, the evaluation brief is likely to serve as an 'agent of change.

Step 5: Determine efficient forms and dissemination methods per evaluation knowledge product

Most evaluation reports and knowledge products can be shared as an electronic copy. In order to enhance the efficiency in terms of time and cost, the organization's public webpage and the e-mail list should be strategically used as means for dissemination. For example, the evaluation reports should be uploaded on the organization's internal and external webpage with a blurb that summarizes the key information in the report. Additionally, knowledge from monitoring and evaluation can be shared widely by incorporating them in existing reports and

publications, such as the country office's annual report or other key reports, brochures and news bulletins. In many cases radio might be an appropriate way of dissemination and written products will simply not comply with current and local conditions. The form of dissemination will be completely different then but the importance of dissemination is not less important.

Step 6: Monitor feedback and measure results of dissemination efforts

There should be a feedback and learning mechanism for the effectiveness of the dissemination strategy and quality of the particular knowledge product. For example, the organizers conduct a quick survey among the recipients of the knowledge products. In analysing the feedback, the following should be asked: "To what extent has the monitoring and evaluation information been used in programming and policy making?"; "Has such information been made in a timely manner to effectively influence decision-making processes?"; "Have the products reached both direct and indirect audiences in an efficient manner and were they easily accessible?"; "Did the audience find the knowledge products useful?"; "If not, why not?"; and "What could be done better next time?" Lessons from the experience should be reflected in the future evaluation knowledge sharing and dissemination efforts so that monitoring and evaluations will continue to be relevant and contribute to organizational learning and the enhancement of a global knowledge base in development.

In detail you will learn more in the next chapters on practice and limits setting up and using a management system

3. Reporting on the information for monitoring and evaluation

Performance information is to be used as a management tool. Thus, performance information is derived from both monitoring and evaluation. Both can provide critical, continuous, and real-time feedback on the progress of a given project, program, or policy. Analyzing and reporting performance findings are critical steps because it determines what is reported, when it is reported, and to whom it is reported. This step also has to address the current technical capacity of the organization because it focuses on the methodological dimensions of accumulating, assessing, and preparing analyses and reports. This chapter focuses specifically on reporting findings and addressing the following issues

Uses of monitoring and evaluation findings;

Knowing the audiences and targeting the appropriate information to those audiences;

Presentation of performance data in clear and understandable form; and

What happens if performance news is bad?

3.1. The Uses of Monitoring and Evaluation Findings

Monitoring and evaluation reports can play many different roles, and the information produced can be put to very different uses:

To demonstrate accountability—delivering on political promises made to citizenry and other stakeholders

To convince—using evidence from findings

To educate—reporting findings to help organizational learning

To explore and investigate—seeing what works, what does not, and why

To document—recording and creating an institutional memory

To involve—engaging stakeholders through a participatory process

To gain support—demonstrating results to help gain support among stakeholders

To promote understanding—reporting results to enhance understanding of projects, programs, and policies. Evaluation reports serve many purposes. The central purpose, however, is to "deliver the message"—inform the appropriate audiences about the findings and conclusions resulting from the collection, analysis, and interpretation of evaluation information.

3.2.   Know and Target the Audience

Know your audiences and how they want to see the information expressed. The interests, expectations, and preferred communications medium of the audience should be taken into account. A communications strategy should be developed that will address the following questions:

Who will receive what information?

In what format?

When?

Who will prepare the information?

Who will deliver the information?

During the ongoing process of determining monitoring and evaluation findings, it is important to ensure that everyone is informed of progress, and that there are no surprises. If the information system is to provide continuous performance feedback as a management tool, continuous communication is also important to the process. Monitoring and evaluation results should be continuously disseminated to provide feedback to decision makers. Informal (phone, e-mail, fax, conversations) and formal (briefings, presentations, written reports) communications should be a part of the overall communications strategy.

Data should be presented in a short and crisp manner and be relevant to the target audience. Only the most important data should be presented.

If there are multiple audiences—those involved at the project, program, and policy levels— the data may have to be packaged and formatted differently according to the main interests

and preferences of each audience. The communications strategy should take into account the challenges in communicating results to different stakeholders. Furthermore: Clear the report with all key parties before it is formally presented. This will help to eliminate errors and will also ensure that many points are clarified informally without the embarrassment of confrontations later on

One can anticipate that there may be multiple uses of the performance findings.

Think of this as concentric circles, that is, the target audience forms the inner circle, but there may be uses for the findings beyond the inner circle including those less directly concerned or affected. Consequently, one should also anticipate further dissemination of performance findings to a broader audience. Typically, the higher up the chain of command, the less need there is for extensive detail and explanation; aggregated, succinct data relevant to the specific issue will be more appropriate. For this reason, personal briefings—especially to high-level officials—can be another effective means of communicating performance findings. Further down the managerial chain, it is more likely that more operational data will be desired. Large "data dumps" of information are counterproductive. Know what the decision makers want and provide them with the necessary information in the format with which they are most comfortable. This may require tailoring information into the preferred format for each of the decision makers and end users.

Automatic dissemination,(websites and others) is not necessary the best way to inform people. There must be a culture of use of using this type of knowledge product and it must respond to all the necessities of decision makers. A good example of a well intended but large unsuccessful knowledge product is DevInfo[9], a perfect tool to disseminate statistical knowledge, but with serious problems using it on a wide scale and as up to date information source for decision making, why:

---

[9] The author would like to avoid generalizing the critical view of DevInfo. The remark is related to detailed practical experiences in several African countries notably Mozambique

people do no understand it, despite its ease of use some introduction is needed to use the software

data are not current,, delay in feeding current data is fatal, people finding only old data will not come again

Access, off-line (on a CD) access is counter-indicative, data cannot be current and a CD can only be used for a limited time

Decision makers may be looking for some indications of action required in response to data findings. They will also be interested in available options (including costs, pros and cons, and the like) with respect to acting on performance findings throughout the monitoring and evaluation process.

Furthermore, it is important to highlight the implications of recommended actions throughout the monitoring and evaluation process. "Simply recommending that certain actions be taken is rarely sufficient. Wise evaluators anticipate this need and provide, whenever possible, best estimates (or perhaps a range of estimates) of both the costs and consequences of the recommendations.

Continuous reporting on findings can and should also extend to guiding decision makers through implementation of recommendations. In terms of follow-up and feedback, one could set up a political process to bring stakeholders and evaluators together to discuss findings, insights, alternative actions, and next steps.

It would also be useful to obtain feedback periodically from major constituencies, such as elected officials, investors, and the public, regarding the usefulness and readability of performance reports. Use the feedback to help tailor future performance reports to the particular audience.

Comparisons of performance data over time are critical. Providing data for a specific quarter or year by itself is not useful. To distinguish trends, one needs to begin with baselines.

Always report against the baseline and intermediate measurements to determine whether progress has been sustained, whether there was only a short spurt of improvement, or whether early improvements have all disappeared. Comparing actual outcomes to targets is central to reporting results. Table 8-1 illustrates indicator baseline and current and target measurements, as well as percentage differences relative to the province mean in bar graphs.

3.3.   **Presentation of Performance Data in Clear and Understandable Form**

It is important to report results data in comparison to earlier data and to the baseline. Comparisons over time are critical. As an example: the following data can be reported:

Expenditure or income—cost of, or return on, project, program or policy • Raw numbers—early indications, rough projections, estimates, and so forth

Percentages (for example, percentage of citizens served by a project)

Statistical tests

Organizational units

Geographical locations

Demographics

Client satisfaction scales—high, medium, low.

Data should be presented in a simple, clear, and easily understandable format. Only the most important data should be presented. Acronyms and jargon should be avoided. A minimum of background information should be provided to establish the context. Major points should be stated up front. Findings and recommendations should be organized around key outcomes and their indicators. A separate appendix or report can be used to convey detailed data. There are four dimensions of reporting: written summaries, executive summaries, oral presentations, and visual presentations.

3.3.1. Written Summaries

To be a useful management tool, the written summary should contain an introduction (including purpose of report, evaluation questions, program background, and program goals and objectives). The summary should contain a description of the evaluation (including evaluation focus, methodology, limitations of methodology, who performed the evaluation, and when the evaluation was performed). The report should present data on findings, selectively and in an understandable manner; organize data around study questions, major themes or program components; and use charts and tables. Conclusions should be clearly connected to evidence on performance. Evidence should be presented to support recommendations. When planning the time needed to prepare the analysis and reporting format, leave plenty of time to revise. Having a knowledgeable outside reader review the findings and draft report can also be helpful.

3.3.2. Executive Summaries

Executive summaries should be short (one to four pages). Major findings and recommendations should be presented in bullet format. The summary can refer readers to the report or appendices for more details. The executive summary should contain a brief overview, including the background and purpose of the study. It should also include a brief description of major questions, issues, and research methods.

3.3.3. Oral Presentations

Oral presentations also can be used, either alone or in conjunction with a written report. In addition to rehearsing and getting feedback, one needs to consider the following in preparing for an oral presentation:

• Who is the audience?

• What should they remember from the presentation?

• How much time is there for the presentation?

• What are the available delivery resources?

• What handouts should be provided, if any?

Oral presentations—like written ones—should be simple, clear, and tailored to the audience. Complex language and detailed data should be avoided. Organization is also important: "Tell them what you will tell them; tell them; tell them what you told them." If possible, use an interactive format with the audience, and be prepared for questions.

3.3.4. Visual Presentations

Visual presentations—charts, graphs, and maps—are also helpful in highlighting key points and performance findings. They can illustrate directions and trends at a glance. There are a variety of charts (pie, flow, column, time series, scatter plot, bar, range, and so forth) and graphs (line, scatter, bar, pie, surface, pictograph, contour, histogram, area, circle, column) that should be considered in presenting data to the target audience. The purpose of charts and tables is to describe, explore, tabulate, and compare. Charts and tables can provide impact and visual interest, encourage audience acceptance and memory retention, and show the big picture. Charts and tables should present data simply and accurately, and make the data coherent. They should engage the audience. Tables are best used for presenting data, and highlighting changes, comparisons, and relationships. Charts are better for presenting the message. They are useful in depicting organizational structures, demonstrating flows, presenting data as symbols, conveying concepts and ideas, and presenting numerical data in visual form. Effectively designed tables will have the following characteristics:

Simplicity and accuracy

Clearly labelled rows and columns with no abbreviations

Percentages rounded to the nearest whole number

Total numbers

Source of the data.

Effective use of white space

Simple

Honest scales

Message conveyed in title

Sufficient data provided with chart so that message is clear

Source of the data

Supporting data in an appendix.

Effective charts enable policymakers and decision makers to quickly see the current status of a given project, program, or policy— including trends, directions, delays, problems, successes, and prospects. Charts should be used to

2012 PRESIDENTIAL RUN
GOP CANDIDATES

BACK PALIN
70%

63%

60%

BACK HUCKABEE   BACK ROMNEY

FOX
47°   SOURCE:OPINIONS
DYNAMIC
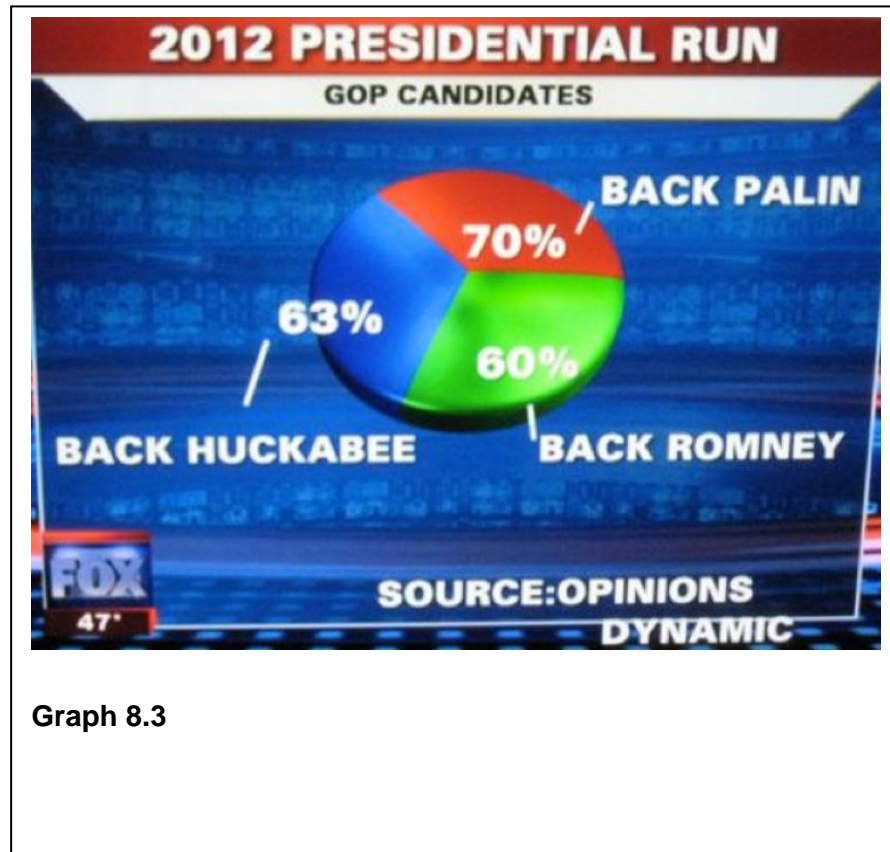
Graph 8.3

provide informative and useful visual aids for continuous reporting of findings. Whether in chart or table form, portraying information graphically is an important part of reporting. Be cautious not to use inappropriate graphs just because they may be popular and you will certainly be careful not to publish a downright ridiculous graph as in 8.3

3.4.   What Happens If the M&E System Produces Bad Performance News?

One cannot manage by receiving only good news. A good performance measurement system is intended to surface problems—not just bring good news. This is another of the political aspects of results-based M&E systems. Reporting on bad news is a critical aspect of how one distinguishes success from failure. If the difference cannot be determined, it is likely

that both failure and success are being rewarded by managers. A good performance system can serve as a kind of early warning system. Performance reports should include explanations (if possible) about poor outcomes and identify steps taken or planned to correct problems. Messengers should not be punished for delivering bad news. Instilling fear of bringing forth bad news will not encourage reporting and use of findings.

4. Using monitoring information

After examining effective ways of reporting in the previous chapter, we turn now to the use of

findings emanating from the results-based monitoring and evaluation system. We will

consider

The uses of performance findings;

Additional benefits of using the findings—feedback, knowledge, and learning; and

Strategies for sharing information.

4.1.   Uses of Performance Findings

Using findings to improve performance is the main purpose of building a results-based M&E

system. The main point of the M&E system is not simply to generate continuous results-

based information, but to get that information to the appropriate users in a timely fashion so

that   the

Ten Uses of Results Findings

1. Respond to elected officials' and the public's demands for accountability
2. Help formulate and justify budget requests
3. Help make operational resource allocation decisions
4. Trigger in-depth examinations of what performance problems exist and what corrections are needed
5. Help motivate personnel to continue making program improvements
6. Formulate and monitor the performance of contractors and grantees
7. Provide data for special, in-depth program evaluations
8. Help provide services more efficiently
9. Support strategic and other long-term planning efforts (by providing baseline information and later tracking progress)
10. Communicate better with the public to build public trust.

Source [17]

performance feedback can be used to better manage organizations and governments.

Findings can be used in a variety of concrete ways, as shown in the following box .

With respect to helping formulate and justify budget requests, performance information can inform decisions that can lead to budgetary increases—or reductions. Projects, programs, and policies may be enhanced or expanded based on performance feedback; likewise, they may be cut or eliminated altogether. Managers also have the option of offering incentives (monetary and non- monetary) to personnel for good performance or sanctions (such as poor employee or manager performance reviews) for performance that fails to meet expectations or falls short of intended outcomes. In terms of motivating personnel, when civil servants are brought in as partners to the business of government, we see better implementation. Employees throughout the system begin to understand and become more enthusiastic about their contributions toward achievement of the desired goal when they have a "line of sight" between their own actions and the goal. In some OECD countries (Australia and France, for

example), managers are given greater operational flexibility in exchange for enhanced accountability. Australia provides an example regarding the performance of contractors and grantees. In Australia, there are actual performance contracts with agencies that specify that no annual budget funds will be allocated until contracts have been evaluated and results monitored.

If there are no data on which to base decisions, those decisions can be arbitrary. At the same time, decision makers always have the discretion to make their own decisions. However, better decision making will result from taking the time to monitor, measure, and evaluate, and incorporate the findings into the decision making process. An interesting corollary to this is that if one starts to ask for performance information, improved performance will result. Other uses of results findings include identifying best practices, supporting economies of scale, avoiding overlap and duplication, and coordinating similar programs across agencies. There are many examples of using findings. The previous box illustrate some of the different uses of performance findings.

4.2. Additional Benefits of Using Findings: Feedback, Knowledge, and Learning

M&E systems provide important feedback about the progress, as well as the success or failure, of projects, programs, and policies throughout their respective cycles. These systems constitute a powerful, continuous public management tool that decision makers can use to improve performance, and demonstrate accountability and transparency with respect to results. The use of M&E findings can promote knowledge and learning in governments and organizations. The new emphasis in the international aid community is more and more on local knowledge acquisition, not knowledge transfer from donor to recipient. What exactly do we mean by "learning" in a results-based monitoring and evaluation context? Learning has been described as a continuous dynamic process of investigation where the key elements are experience, knowledge, access and relevance. Knowledge and knowledge management are additional key components of using performance findings. New knowledge can be generated through the use of findings on a continuous basis. Knowledge management

means capturing findings, institutionalizing learning, and organizing the wealth of information produced continually by the M&E system. Results-based monitoring and evaluation systems and units have a special capacity to add to the learning and knowledge process. When used effectively, M&E systems can be an institutionalized form of learning and knowledge.

> ### Obstacles to Learning
>
> The OECD has identified several obstacles that can prevent learning:
>
> *Organisational culture*—some organisations have a culture where accountability tends to be associated with blame. This has the effect of discouraging openness and learning. In other [organizations], it is more acceptable to own up to mistakes and see these as opportunities for learning, recognizing that there is often as much to learn from poorly performing projects as there is from success stories.
>
> *Pressure to spend*—learning takes time, and pressure to meet disbursement targets can lead to shortcuts being taken during project planning and approval stages, with lessons from previous experience being ignored or only selectively applied in the haste to get decisions through.
>
> *Lack of incentives to learn*—unless there is proper accountability . . . built into the project cycle there may be little incentive to learn. This is particularly the case when staff or consultants shift from task to task, and have generally moved on long before the consequences of failure to learn are felt.
>
> *Tunnel vision*—the tendency of some staff or operational units to get stuck in a rut, carrying on with what they know, even when the shortcomings of the old familiar approaches are widely accepted.
>
> *Loss of institutional memory*—caused by frequent staff rotation or heavy reliance on short-term consultants, or by the weakening or disbanding of specialist departments.
>
> *Insecurity and the pace of change*—if staff are insecure or unclear about what their objectives are, or if the departmental priorities are frequently shifting, this can have an adverse effect on learning.
>
> *The unequal nature of the aid relationship*—which tends to put donors in the driving seat, thereby inhibiting real partnerships and two-way knowledge sharing.
>
> [19] OECD 2001, pp. 20–21.

Institutionalizing learning is important in governments and organizations. Policy and program evaluation should play a systematic instead of an ad hoc role in the process of organizational learning. A political environment needs to be created that encourages continuous reporting, as well as the use of results. This implies that a certain level of institutionalization has to occur before findings can be used in the management of government institutions. Emphasizing organizational learning as a means of enhancing organizational performance is a fruitful and promising area of engagement with the public sector. Many governments and organizations may yet be resistant to learning, internalizing, and sharing performance findings within and between ministries, organizations, agencies, and departments. There are a number of organizational, behavioural, and political challenges to be recognized. Good M&E systems can help to overcome these obstacles to learning. By producing a continual flow of feedback and data, M&E systems help decision makers manage more effectively. Organizational cultures can be transformed through the use of M&E systems. There may be decreased pressures to spend as governments receive data that help them manage resource flows. M&E systems also provide built-in incentives to learn, pointing out directions, trends, successes, and problems. Tunnel vision can be overcome as data on results shed light on areas previously unknown or not fully understood.

The loss of institutional memory due to staff changes can also be minimized because M&E systems, when well maintained, produce a record of data over time. Finally, change can be managed more easily with continuous feedback. Obstacles can also be overcome by understanding how governments and organizations learn and by identifying and overcoming the impediments. There are ways to encourage greater use of performance findings through learning and knowledge building among governments and organizations.

4.3.   Strategies for Sharing Information

A good communication strategy is essential for disseminating information and sharing it with key stakeholders. Results-based information should be shared with all internal and external stakeholders and interested parties. Information sharing strategies designed for and targeted to specific stakeholder groups can also be helpful. Using results information can take passive and active forms. Understanding the target audience is the key. Communication strategies need to be tailored to suit a particular target audience—parliament, ministers, the media, the private sector, NGOs and civil society organizations, and the general public. Governments and organizations can use a wide array of strategies for sharing information with internal and external stakeholders. These strategies also involve a number of different media that can be used to share the performance information.

Empower the Media.

The media can be an important partner in disseminating the findings generated by results-based M&E systems. For example, the media often report on whether governments or organizations have actually delivered on promised projects, programs, policies, and services. The media have also been instrumental in exposing corruption and calling for good or better governance in many countries.

Enact "Freedom of Information" Legislation

Freedom of information is another powerful tool that can be used to share information with concerned stakeholders. For example, the government of Romania enacted freedom of information legislation recently with the stipulation that, except for information that could

impair the country's ability to protect and defend itself, anyone who asks for information about how well the government is performing will receive it.

Institute E-Government

E-Government is increasingly being used as a tool by governments around the world, and has become a particular priority among OECD countries. E-government involves the use of information technology to provide better accessibility, outreach, information, and services. It represents a new electronic environment in which stakeholders can interact directly with the government, obtain information from the government, and even transact business online. Developing countries are moving in this direction, too.

Put Information on Internal and External Internet Sites

The use of internal (agency or government) and external Web sites that include published performance findings is yet another effective way of sharing information. Many agencies are also developing searchable databases for M&E findings.

Publish Annual Budget Reports

There is no more important way to communicate how taxpayer money is being spent than to publish the budget. Citizens will have the opportunity to "compare" the quality and level of services being provided by the government, and the priority of that service or program in the expenditure plan.

Engage Civil Society and Citizen Groups

Engaging civil society and citizens groups also involves the inclusion of accountability, advocacy and action-oriented audiences and agreement on the information (content and form) they need.

Strengthen Parliamentary Oversight

Strengthening parliamentary oversight is another important way to share and disseminate information. Many parliaments have active budget or public accounts committees in lower or upper chambers. Parliaments in various countries—both developed and developing—are starting to ask for performance information as part of their oversight function. They are

looking to see that budgets are used effectively; thus, more governments are considering moving toward programmatic budgeting.

Share and Compare Results Findings with Development Partners

Sharing and comparing results findings with development partners is also beneficial on a number of levels. Since the introduction of National Poverty Reduction Strategies and similar broadly based strategies and policies, the need for information sharing among development partners—especially bilateral and multilateral aid agencies—has increased. "These and other joint initiatives are premised on the assumption that coordinated agency action will be more effective than individual efforts. Yet mechanisms for exchanging evaluation lessons between [aid] agencies are still weak, and practical hurdles continue to get in the way of more frequent joint evaluations—which, when they do occur, are generally seen as a very good way of sharing lessons and methodologies" [19]. More could also be done with respect to sharing performance findings with donor recipient countries. All key stakeholders— particularly recipient countries—need to be part of the M&E process from start to finish.

There are many uses for performance findings. We looked at two successful examples involving crime information and a government organization with a mature, functioning M&E system. We also examined the many benefits of using findings, including continuous feedback, and organizational and institutional learning and knowledge. We acknowledged and examined the obstacles and incentives— many of them political—to using findings, and looked at some potential strategies for sharing information among internal and external stakeholders.

5. Sustaining the management system

The purpose of a management system is not only to monitor but essentially to promote good decision making. So in this context we will speak of a management system to bridge the gap from monitoring to evaluation and onwards to decision making. Here, we turn to sustaining results-based M&E systems and we favour the word Management system instead because it implies much more the decision making and political impact of such a system. In fact a whole policy dialogue should be grouped around this issue which is far beyond a classical technical tool. A Management system should be regarded as a long-term effort, as opposed to an episodic effort for a short period or for the duration of a specific project, program, or policy. Sustaining such systems within governments or organizations recognizes the long-term process involved in ensuring utility (for without utility, there is no logic for having such a system). Specifically, we will examine:

Six critical components of sustaining results-based Management systems;

The importance of incentives and disincentives in sustaining Management systems;

Possible hurdles in sustaining a results-based Management system ;

Validating and evaluating Management systems and information; and

M&E stimulating positive cultural change in governments and organizations.

5.1. Six Critical Components of Sustaining Results-Based Management systems

We will examine six critical components involved in building the sustainability of Management systems. Each of these dimensions needs continuous attention and care.

Demand

If demand is episodic or haphazard, results-based Management systems are not going to be used and sustained. Structured requirements for reporting results, including legislation, regulations, and international development requirements, can help lead to sustained, consistent demand for such systems. Governments, civil society, and donors are increasingly requiring the results that Management systems can best track, monitor, and measure. In many cases, demand can also be stimulated when the strategic goals of the government are translated into results-based Management systems, such as through National Poverty Reduction Strategies and other initiatives. These are not simply activity-driven initiatives; rather, they try to answer the "so what" question. What are the consequences of policy and program efforts to reduce poverty and address the most vulnerable groups?

Examples: Reports on Migration, Poverty and Political Unrest will keep awareness high in countries otherwise far way from the centres of unrest.

Clear Roles and Responsibilities

Clear roles and responsibilities and formal organizational and political lines of authority must be established. The organization and people who will be in charge of collecting, analyzing, and reporting performance information must be clearly defined. Guidance is necessary. For example, a Ministry of Finance may be responsible for administering National Poverty Reduction Strategies or initiatives, and will need to issue directions to the sector or line ministries to collect and report on data relevant to tracking the various outcomes specified in the strategy.

Internal political coordination is a key. A system should be built that links the central planning and finance ministries to the line and sector ministries. These bridges linking ministries are important, as is the need for horizontal communication to keep all concerned parties informed. If there are organizational problems, these should be dealt with sooner rather than later. It is also important to build a continuous system of data collection and analysis that goes beyond the national government to other levels of government.

Data collection, analysis, and reporting should be aligned throughout the various levels of government. For example, in the health or education sectors, focusing at the local and regional levels will be important because some of the requirements to meet national goals are going to take place there. Data analysis and reporting at these levels will then feed into the larger national data base in determining progress toward the desired outcomes. Finally, Management systems should be built in such a way that there is a demand for results information at every level that data are collected and analyzed. There is no level of the system that is a mere "pass through" of information. Pass-through parts of the system create tremendous vulnerability, and can lead to breakdowns in Management systems. If people are not involved, if there is no ownership, then people in the "pass-through" levels will begin to lose interest and the result will be poor data collection and reporting.

Trustworthy and Credible Information

The Management system must be able to produce results information that brings both good and bad news. Performance information should be transparent and made available to all key stakeholders. If debate of issues is not backed up by trustworthy and credible information, only personal opinions and presumptions are left. It should also be noted that the producers of results information need protection from political reprisals. If bad news brings career problems to the messengers, fear will permeate the system and the reliability of the information produced will be compromised.

A quick way to undermine a Management system is to punish those who deliver bad news. Information produced by the Management system should be transparent and subject to independent verification. If data on government performance are held too close, or there are gatekeepers who prevent the release of such information, the system will again be faulty. As a further check on the system, it would be advisable to have a periodic independent review by the national audit office, parliament, or a group of academics to ensure that the data being generated by the system are accurate and reliable, and to build confidence among managers who could use the data.

Accountability

No part of the government should be exempt from accountability to stakeholders. Civil society organizations and NGOs (such as Transparency International) can play a key role in encouraging transparency and accountability, and can even help with collecting data. For example, NGOs in Bangladesh help to collect local educational data because the capacity to collect and report on such data is very weak within the government. The media, private sector, and parliament also have roles to ensure that the information produced is timely, accurate, available, and addresses government performance. It is also important not to reward failure. Accountability means that problems should be acknowledged and addressed.

Capacity

Sound technical skills in data collection and analysis are necessary for the system's sustainability. Managerial skills in strategic goal setting and organizational development are also needed. Data collection and retrieval systems must be up and running—and modernized. Governments will need to commit continuing financial resources to the upkeep and management of results-based Management systems. Institutional experience and memory are also helpful in the long-term sustainability of these systems.

Incentives

Incentives need to be introduced to encourage use of performance information. This means that success needs to be acknowledged and rewarded, problems need to be addressed, messengers must not be punished, organizational learning is valued, and budget savings are shared. Corrupt or ineffective systems cannot be counted on to produce quality information and analysis. Developing countries are also working toward creation of evaluation capacity, institutionalization of evaluation, and use of results findings within government—in short, sustainable Management systems. Table 8.3 provides a comparative illustration of such efforts in Colombia, China, and Indonesia.

5.2. The Importance of Incentives in Sustaining Management systems

Sustaining Management systems also involves using appropriate incentives to keep managers and stakeholders on track and motivated. "Putting in place incentives for M&E means offering stimuli that encourage . . . M&E officers and primary stakeholders to perceive the usefulness of M&E, not as a bureaucratic task, but as an opportunity to discuss problems openly, reflect critically and criticize constructively in order to learn what changes are needed to enhance impact" (IFAD 2002, Section 7 p. 4). There are a variety of organizational, financial, resource, political, technical assistance, and training incentives that can be used to sustain Management systems. Likewise, managers need to remove disincentives to sustaining Management systems. Boxes 10.3 and 10.4 contain checklists of the kinds of incentives and disincentives that should be considered.

5.3. Possible Problems in Sustaining Results-Based Management systems

There are a number of hurdles that may arise in sustaining Management systems. Hatry (1999) brings to light a number of likely problems in implementing and sustaining Management systems, as follow:

• Personnel training needs

• Overall system cost and feasibility

• Changes in legislative and agency priorities

• Maintaining indicator stability over time

• Documentation of the outcome measurement process (who will do what)

• Fear and resistance from program managers

• Participation by other levels of government and the private sector

• Aggregation of outcomes across projects, programs, or sites

• Community-wide versus program-specific outcomes

• Legislative support

• Politics.

Some of the most critical issues in implementing and sustaining

Management systems are the challenges in the human resource area. These challenges are

perhaps not so different from all public sector human resource matters, but there are unique

dimensions that have to be addressed. First, there are issues in recruiting and holding

talented staff who can build and manage a new information system. Can they be found and,

if so, can they be hired? Second is the issue of what staff will risk venturing into a new

government initiative—or stated differently, what is the caliber of those who leave their

present positions for positions in a new M&E unit? Third is the matter of whether the first

cohort of those hired are change agents. Building an Management system is a politically

charged change process. Do those being hired understand this and are they ready to

manage a change process? Fourth, can continuous training be provided for all personnel at

all levels? New methodologies, technologies, and procedures are inevitable and need to be

shared with staff. Can that training be provided? Furthermore, given staff turnover, how soon

and how adequately can new staff be trained to quickly increase their productivity and

contributions to the unit? The Management system will have to respond and adapt to

changes in legislative and organizational priorities. In spite of these larger political and

environmental changes, maintaining indicator stability over time is important. One wants to

be able to compare similar issues and trends over a given period of time.

5.4.   Validating and Evaluating Management systems and Information

Continued upgrading and improvement is important in sustaining results- based

Management systems. Management systems themselves should be evaluated periodically,

using internal or external evaluators. "Evaluators can assist in validating performance data

and improving performance measurement systems. Evaluations of performance

measurement systems should focus both on the technical quality of the measurement

system and on the extent to which performance information is used in managing to achieve

performance goals and in providing accountability to key stakeholders and the public"

Module 8: – Setting-up and Using a Monitoring System: How to improve decision making

(Wholey 2001, p. 345). Evaluators can also verify and confirm the results of Management systems.

5.5.   M&E: Stimulating Positive Cultural Change

Management systems are essentially political challenges, and to a lesser extent, technical ones. Creating, implementing, and sustaining results-based Management systems can help to bring about major cultural changes in the way governments and organizations operate. Management systems can bring about positive cultural changes that lead to improved performance, enhanced accountability and transparency, and learning and knowledge (see box 10.5). Good results-based Management systems must be used to be sustainable. Six components are necessary in sustaining these systems: demand, incentives, clear roles and responsibilities, trustworthy and credible information, accountability, and capacity. Sustainable Management systems do exist in many OECD countries, and some developing countries are on their way toward building and sustaining such systems as well. Above all, results-based Management systems are powerful public management tools that facilitate positive cultural and political changes in governments and organizations to demonstrate results, accountability, and transparency. They also facilitate knowledge and learning. And, they are doable!

Last Reminders

• The demand for capacity building never ends. The only way an organization can coast is downhill.

• Keep champions on your side and help them.

• Establish the understanding with the Ministry of Finance and the parliament that an Management system needs sustained resources.

5.6. Exercise on Elements and Issues of a Results-Based Management systems

Please fill the empty cells, a roster of issues for a Management System (MS) has been designed with examples from a generic system in Columbia and China (for reference).

Issue	Colombia	China	Your Choice
Anchoring the evaluation regime	Constitution mandates the Executive to take the lead.	State Council draft resolution calls on the Central Executive Agencies to take lead.	
Positioning the evaluation function	Centralized in the National Planning Department (NPD). Key line agencies provide inputs.	Decentralized in key central agencies.	
Evaluation coverage	Public policy and major public sector programs.	Public sector projects.	
Linking evaluation with other public sector functions	NPD plays a key role in policy and strategy formulation and budget allocation and monitoring.	No formal links have been established. State Planning Commission involved in public resources allocation and monitoring.	
Using evaluation in decision making	Monitoring and evaluation information to flow to line agency heads and the NPD.	Monitoring and evaluation to inform central agency management.	
Professionalizing the evaluation function	Evaluation is a trans-discipline cutting across specific professional skills.	Evaluation is seen primarily as applied socioeconomic analysis.	
Resources for evaluation	Evaluation to be main-streamed in agencies' budgets.	Evaluation mainstreamed in central agencies' budgets.	

6. Bridging the gap from Monitoring to Evaluation

6.1. Introduction

Monitoring and evaluation are often described in two contradictory ways. The first is to refer in one breath to "emandee" (M&E) as if they are the same thing and can be dealt with by the same people, using the same data, at the same time. The second is to describe evaluation as the job of experts in evaluation, with no connection to monitoring at all.

This course has concentrated so far on monitoring, since it is generally accepted that it needs more emphasis and it is often done badly. We now have to show how monitoring is connected with evaluation, and how, as we said at the beginning:

if the intended results are not clearly stated, monitoring cannot be done well

if monitoring is badly organised, the necessary data are not collected

if the data are not collected, the programme can neither be assessed nor evaluated.

We shall now consider what comprises an evaluation, and show that it does not have a single definition.

6.2. A modern view of evaluation

Kusek and Rist, in *Ten steps to a results based monitoring and evaluation system (WB 2004) [19]*, stress that the definition of evaluation has changed over the decades. The traditional view of evaluation was something which was done when the project or programme was over. This has the advantage of being (usually) a thorough piece of independent research, which contributed greatly to our detailed understanding of what worked and what did not, and why.

But the reality is that many development projects and programmes are followed in quick succession by a second phase or an extension, and planning for this is taking place as the first phase is still running. It is clearly useful to be able to justify the further programme with more scientific evidence than a "feeling that all is OK".

For this, we need evaluation material as well as monitoring evidence, and we need it preferably while the project is running. It is not sufficient to point to progress and the successes of the first phase. We also need the counterfactuals – what did <u>not</u> work so well, which intended beneficiaries were excluded <u>and why</u> – in short, how a successor project can be made more effective.

6.3.   Different types of evaluation

Kusek and Rist identify seven types of evaluation:

1. Performance logic chain assessment
2. Pre-implementation assessment
3. Process implementation evaluation
4. Rapid appraisal
5. Case study
6. Impact evaluation (the classic "after-the-fact" evaluation
7. Meta-evaluation

Some of these will be reviewed in Unit 9. At this stage it is sufficient to note that many of these various evaluations exist simultaneously with monitoring, and are in fact complementary to it.

6.4.   Complementarity of evaluation and monitoring

Three types of complementarities have been identified:

- Sequential complementarity, in which monitoring information generates questions to be subsequently answered by evaluation, or the reverse, where evaluation information gives rise to new areas or domains to be monitored;

- Information complementarity, where both monitoring and evaluation can use the same data but pose different questions and undertake different analyses; and

- Inter-actional complementarity, where managers use monitoring and evaluation at the same time to help to direct their projects and programmes.

The way in which we involve stakeholders in programme design, choice of indicators and monitoring, has a strong link to whether these various evaluations will be able to contribute fully to a better understanding of what makes a good project.

6.5.   Data issues

Monitoring systems generate much of the data which will be required by those carrying out an evaluation, and evaluation needs should be considered as early as possible when designing the monitoring systems. This is crucial for a number of reasons.

First, it can be more economic and more efficient to collect and record certain information while the programme or project is running, rather than after the event. This includes data on the following:

What criteria were used to select programme beneficiaries (often this is part of the baseline assessment), and how these criteria have changed over time;

How the programme is being implemented, and to what extent this follows the plan;

whether beneficiaries benefit equally from the programme;

How many people drop out of the programme, for what reasons, and what are their characteristics compared with those who remain; and

How programme outputs compare with the original plan.

Second, there is the issue of motivating staff to take pride in accurate monitoring. Too often they do not see how, or even if, their data are being used. It is felt that if evaluation findings are more widely disseminated, then over time this can create an incentive for people to improve the quality of monitoring data.

Third, evaluation is not only carried out at the end of the project. If the evaluation team has access to monitoring data, then they may be able to make interim reports – perhaps on issues related to impact or outcome – which will add value to the basic monitoring reports.

Fourth, much monitoring information is collected through national survey programmes. If is increasingly possible to add questions or modules to existing surveys which can probe in more detail certain outcomes. A survey may note, for instance, that a child in the programme target area is no longer at school. An additional question could try to establish why, when the issue is current, rather than several years later.

Fifth, and perhaps the most important, the monitoring team need to keep the information they are generating in a form which allows for access by the evaluation team, and also the type of analyses they may wish to do. This might mean coding or storing the data slightly differently from classical ways.

6.6.   Summary

On a day to day basis, those involved with implementing a project or programme do not see the whole picture – only a very small part of it. But it is necessary for everyone to have an idea of the whole and their place in it. In particular, everyone needs to know why the money is being spent, the efforts and sacrifices made. This may be called the Goal, or the Vision, or some other all-embracing term – the exact terminology is unimportant – but what matters is that it is shared with all the members of the team.

A clear understanding of the goal helps all stakeholders to place their own contribution in context. This applies to all aspects of implementation. In the context of this course, it means that everyone involved with a project is generating information about progress, in one way or another, and everyone should know how this information can contribute to both monitoring and evaluation.

And those who are designing and supervising the programme monitoring systems need to be particularly aware of the multiple uses for the information generated.

Bibliography

Documents

UNEG, 'Norms for Evaluation in the UN System', 2005, available at:

http://www.unevaluation.org/unegnorms ; and UNEG, 'Standards for Evaluation in the UN

System', 2005, available at: http://www.unevaluation.org/unegstandards.

World Bank Group: Handbook on Impact Evaluation Quantitative Methods and Practices,

Shahidur R. Khandker, Gayatri B. Koolwal, Hussain A. Samad: Handbook Evaluation-WB.pdf

Making Smart Policy: Using Impact Evaluation for Policy Making- World Bank Group June

2009

Has Economic Growth in Mozambique Been Pro-Poor? Discussion Paper 2005 International

Food Policy Research Institute Washington D.C: IFPRI_hasDevelopmentbeenProPoor.pdf

Is Poverty Decreasing in Mozambique? Joseph Hanlon, Senior Lecturer,Open University,

England 2007

Do bicycles equal development in Mozambique? Von Joseph Hanlon,Teresa Smart

Rochester NY and Woodbridge: James Currey, 2008. ISBN 978 184701 319 4.

Project Proposal for a Survey Measuring Development in Sofala on District 2007

LevelKRoeder_ProposingSMD_Sofala070301.pdf

Heterogeneous Impacts of Conditional Cash Transfers: Evidence from Nicaragua Ana C.

Dammert - Dammert_Nicaragua_2008.pdf

Information Distortion, Elite Capture, and Task Complexity in Decentralised Development by

Jean-Philippe Platteau - Information Distortion, Elite Capture.. in Decentralized

Development- Platteau - 2008.pdf

Evaluating three stylised interventions Author: Martin Ravallion – in Development

Effectiveness - Affiliation: a Development Research Group, World Bank, Washington, DC,

USA

ASTA Wirtschafts- und Sozialstatistisches Archiv Band 4-Heft 4-2011, J.Frohn Glück,

Nachhaltigkeit und Ökonometrie Zur Einbeziehung sozialer Aspekte in makroökonomoischer

Modelle

Hatry, Harry P. 1999. Performance Measurement: Getting Results. Washington, D.C.: The

Urban Institute Press.

OECD (Organisation for Economic Co-operation and Development). 2001. "Evaluation

Feedback for Effective Learning and Accountability." Paris: OECD/ DAC.

Ten Steps to a Results- Based Monitoring and Evaluation System, Jody Zall Kusek and Ray

C. Rist, The World Bank 2004

Links

http://www.google.de/url?sa=t&source=web&cd=1&ved=0CB0QFjAA&url=http%3A%2F%2Fs

iteresources.worldbank.org%2FINTPA%2FResources%2FTraining-

Materials%2FMozambique_BeatingTheOdds_handout.pdf&rct=j&q=World%20Bank%2C%20

Beating%20the%20Odds%3A%20Sustaining%20Inclusion%20in%20a%20Growing%20Eco

nomy%20%E2%80%93%20A%20Mozambique%20Poverty%2C%20Gender%20and%20So

cial%20Assessment%202007.&ei=1z1dTZS9IsfasgahwuHeCg&usg=AFQjCNHw9dL9J3NA7

HEqorBFPF1sz1V2VA&sig2=BR2I572_1Qv4NosyAeN7gw

No link available

http://books.google.de/books?id=kVs2QiJof_8C&printsec=frontcover&dq=Joseph+Hanlon+a

nd+Teresa+Smart.+Do+Bicycles+Equal+Development+in+Mozambique&source=bl&ots=V6u

A2nSVlH&sig=vXSDuG5nJbpxrsY0TmZ50dFyM2w&hl=de&ei=nkxdTaLINYKi8QPhgb2KCw

&sa=X&oi=book_result&ct=result&resnum=1&sqi=2&ved=0CBgQ6AEwAA#v=onepage&q&f

=false

http://www.klaus-roeder.com/4_Projekte/Trainings_Projects/1987_1990%20Harambee.html

ftp://ftp.fao.org/docrep/fao/010/i0132e/i0132e03b.pdf

http://lnweb90.worldbank.org/oed/oeddoclib.nsf/DocUNIDViewForJavaSearch/E629534B7C6

77EA78525754700715CB8/$file/inst_ie_framework_me.pdf

Better Monitoring for Better Evaluation

Module 9: Evaluation

Content

1. Introduction and Preparation

1.1.   Aim and Learning goals

The part should answer the question: What if we monitored all the development? What will be ways to make a concise sand understandable evaluation at the different points and milestones of the monitored process or project? Maybe it will raise more questions because there is no easy and simple answer to this. The earliest definitions of the monitoring framework will find its way to this part as well because it is not possible to conceive any evaluation without a related framework and assigned roles and responsibilities. The set of three modules should wrap up the final goals of a system cumulating in Better Monitoring and better Evaluation for better decision making and politics.

After completing this module, the reader or participant will be able to discuss and understand

Different approaches to program evaluation

Differences between quantitative and qualitative approaches to evaluation,

Ways selection bias in participation can confound the treatment effect

Different methodologies in impact evaluation

Different examples (best practices of evaluation)

The ultimate learning effect will be that reader or participant will be able to set up a flexible and efficient system to evaluate progress and failure

1.2.   Initial Words / Abstract

Good planning, monitoring and evaluation enhance the contribution of any manager or planner by establishing clear links between past, present and future initiatives and development results. Monitoring and evaluation can help an organization to extract relevant information from past and ongoing activities that can be used as the basis for programmatic

fine-tuning, reorientation and future planning. Without effective planning, monitoring and evaluation, it would be impossible to judge if work is going in the right direction, whether progress and success can be claimed, and how future efforts might be improved. This chapter describes the purposes of planning, monitoring and evaluation in the context of results-based management (RBM) and managing for development results (MfDR) and explains how these functions are important. It also provides key definitions and principles that are integral to planning, monitoring and evaluation. This chapter is intended for managers, staff, key partners and stakeholders who are involved in the design and implementation of development initiatives and decision making. It is advocated here that the culture of results orientation should be embraced by all in order to effectively contribute to human development.

MfDR is RBM in action, but it is oriented more towards the external environment and results that are important to programme countries and less towards an agency's internal performance.

Achieving development results, as most realize, is often much more difficult than imagined. To achieve development results and changes in the quality of people's lives, governments, development partners will often develop a number of different plans, strategies, programmes and projects. These typically include:

A National Development Plan or Poverty Reduction Strategy

Sector-based development plans

A United Nations Development Assistance Framework (UNDAF)

A corporate strategic plan (such as the UNDP 2008-2011 Strategic Plan)

Global, regional and country programme documents (CPDs) and country programme action plans (CPAPs)

Monitoring and evaluation (M&E) frameworks and evaluation plans

Development and management work plans

Office and unit specific plans

Project documents and annual work plans

However, good intentions, large programmes and projects, and lots of financial resources are not enough to ensure that development results will be achieved. The quality of those plans, programmes and projects, and how well resources are used, are also critical factors for success.

To improve the chances of success, attention needs to be placed on some of the common areas of weakness in programmes and projects. Four main areas for focus are identified consistently:

Planning and programme and project definition

Stakeholder involvement.

Communication.

Without proper planning and clear articulation of intended results, it is not clear what should be monitored and how; hence monitoring cannot be done well.

Without effective planning (clear results frameworks), the basis for evaluation is weak; hence evaluation cannot be done well.

Without careful monitoring, the necessary data is not collected; hence evaluation cannot be done well.

Monitoring is necessary, but not sufficient, for evaluation.

Monitoring facilitates evaluation, but evaluation uses additional new data collection and

---

Monitoring and evaluation

Inter-linkages and dependencies between planning, monitoring and evaluation

1.3.   About this Module

This module presents a holistic view of evaluation in order to help managers and staff of programme units and partners make strategic decisions about evaluation.

The chapter describes why evaluation is important for and how evaluative information should be used, then briefly presents the evaluation policy, types of evaluations that are commonly conducted in , key roles and responsibilities in evaluation, and evaluation requirements as stipulated in the evaluation policy.

It will finish with some practical examples. The purpose is a learning process for best practices; a training course will give room for practice and feedback.

1.4.   *Evaluation, what is it*
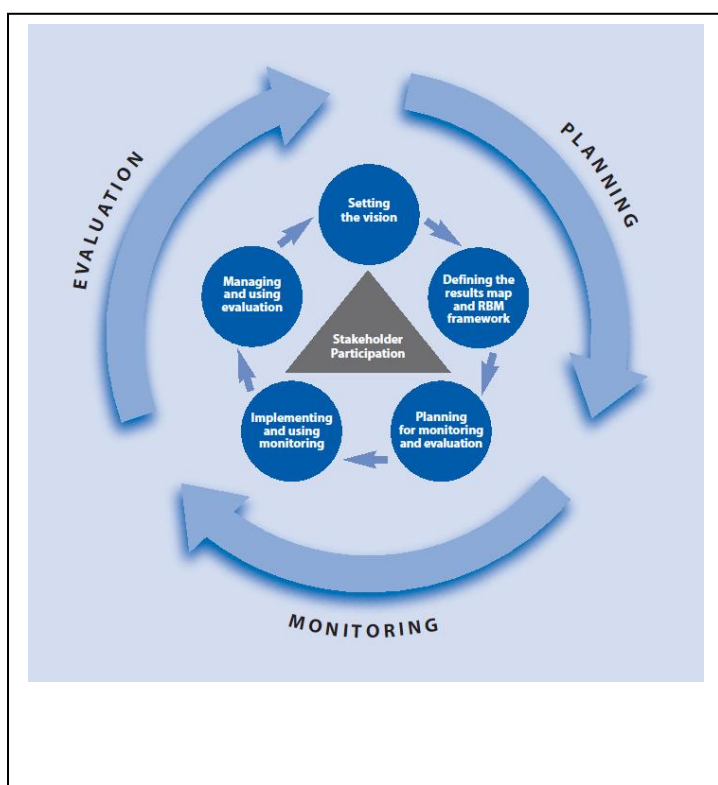
Evaluation is a rigorous and independent assessment of either completed or ongoing activities to determine the extent to which they are achieving stated objectives and contributing to decision making. Evaluations, like monitoring, can apply to many things, including an activity, project, programme, strategy, policy, topic, theme, sector or organization. The key distinction between the two is that evaluations are done independently to provide managers and staff with an objective assessment of whether or not they are on track. They are also more rigorous in their procedures, design and methodology, and generally involve more extensive analysis. However, the aims of both monitoring and evaluation are very similar: to provide information that can help inform decisions, improve performance and achieve planned results.

However it is clear, that Monitoring and Evaluation, often used in conjunction are very different and have distinct tools and procedures albeit joint objective. Planning, monitoring

and evaluation should not necessarily be approached in a sequential manner. The conduct of an evaluation does not always take place at the end of the cycle. Evaluations can take place at any point in time during the programming cycle. The following figure aims to illustrate the inter-connected nature of planning monitoring and evaluation

1.5.   Specific Considerations for Planning Evaluations

It is mandatory for UNDP to present an evaluation plan to its Executive Board with each country, regional and global programme document considered for approval. The evaluation plan is a component of the M&E framework and should include those evaluations that can be foreseen at the end of the programme planning stage. The plan should be strategic, including a selection of evaluations that will generate the most critical and useful information for UNDP and its partners in decision making. The initial evaluation plan should, at a minimum, include **all mandatory evaluations**. For programme units in UNDP, outcome evaluations and project evaluations required by partnership protocols such as the Global Environment Facility are mandatory. The evaluation plan is not a static document. It should be reviewed as part of the M&E framework and refined as needed during programme implementation. For example, as new projects are designed and the needs for evaluations are identified, these new evaluations should be added to the evaluation plan. After a country, regional or global programme is approved, the respective programme unit enters the evaluation plan in the Evaluation Resource Centre (ERC) for tracking.25 As the units exercising oversight responsibility, the regional bureaux use the evaluation plan submitted by the programme units as the basis for assessing compliance. The Evaluation Office reports on evaluation compliance directly to the UNDP Executive Board in its Annual Report on Evaluation. UNDP programme units are required to select and commission evaluations that provide substantive information for decision making. In deciding what to evaluate, the programme units should first determine the purpose of the evaluation and other factors that may influence the relevance and use of proposed evaluations. In general, for accountability purposes, at least

20 percent to 30 percent of the entire programme portfolio should be subject to evaluation. Evaluations generally require significant resources and time. Therefore, every evaluation must be justified and used in an optimal way. Programme units together with key stakeholders should consider the following points in developing an evaluation plan:

1.5.1. Uses, purpose and timing of evaluation

Evaluations should be proposed only when programme units and stakeholders are clear at the onset about why the evaluation is being conducted (the purpose), what the information needs are (demand for information), who will use the information, and how the information will be used. Such information can be derived from a shared vision of success, as expressed in the results or outcome model at the planning stage. The intended use determines the timing of an evaluation, its methodological framework, and level and nature of stakeholder participation. The timing of an evaluation should be directly linked to its purpose and use. To ensure the relevance of an evaluation and effective use of evaluation information, the evaluation should be made available in a timely manner so that decisions can be made informed by evaluative evidence.

1.5.2. Resources invested

An area (thematic or programmatic area, outcome or project) in which stakeholders have invested significant resources may be subject to an evaluation as there may be greater accountability requirements.

1.5.3. The likelihood of future initiatives in the same area

Evaluations are an important means of generating recommendations to guide future work. An evaluation enables the programme unit to take stock of whether the outputs have contributed to the outcome and whether UNDP has crafted an effective partnership strategy. When

selecting an initiative to be evaluated, look for one in an area that UNDP will continue to support.

1.5.4. Anticipated problems

Evaluations can help prevent problems and provide an independent perspective on existing problems. When selecting an outcome for evaluation, look for those with problems or where complications are likely to arise because the outcome is within a sensitive area with a number of partners.

1.5.5. Need for lessons learned

What kinds of lessons are needed to help guide activities in this country or other countries or regions in the region?

1.5.6. Alignment and harmonization

Planned evaluations should be aligned with national, regional and global development priorities and corporate priorities (for example, a national Strategic Plan), and should be harmonized with evaluations of UN system organizations and other international partners. This ensures that proposed evaluations will generate important information to help UNDP and its partners better manage for results in a changing context. Opportunities for joint evaluations with governments and partners should be actively pursued. Evaluations commissioned should be useful for national partners.

In determining the timing of an evaluation, one should consider various decision-making points that exist in the partner government, such as budget decision making, development framework or strategy setting, and existing review processes for development programmes and projects. For instance, if the government is undertaking an evaluation of a national development strategy or framework to which UNDP projects are contributing, the UNDP-

managed evaluations should enhance complementarities and minimize duplicated efforts. Once the outcome evaluations are selected, the programme unit identifies the projects that are designed to contribute to the outcome and indicates them as relevant projects for the evaluation plan. This gives notice to the concerned projects and allows them to take account of the outcome evaluation in their monitoring and work planning. It also helps the UNDP programme officers and relevant national partners in outcome monitoring prepare for the outcome evaluation. The same criteria for selecting outcomes should be applied to selecting **project evaluations**. Some partnership protocols require their related projects to be evaluated.

It is strongly recommended that evaluations should be completed for pilot projects before replication or upscaling, projects that are going into a next phase, and projects ongoing for more than five years for accountability and learning purposes. As part of the regular updating process of the evaluation plan, any newly identified project evaluations should be included in the plan. In crisis settings, extra time should be allocated to evaluations, as there is a need for flexibility in order to respond to changing situations. This means being flexible when scheduling field visits and interviews and anticipating delays in data collection and last-minute changes in data collection methods if relationships between different groups change. Further, more preparation is required when working with vulnerable groups and those affected by conflict, as greater care and ethical considerations are required.

1.6.   Results-Based Management Joint Planning, Monitoring and Evaluation

Planning, monitoring and evaluation come together as RBM. RBM is defined as "a broad management strategy aimed at achieving improved performance and demonstrable results," and has been adopted by many multilateral development organizations, bilateral development agencies and public administrations throughout the world (as noted earlier, some of these organizations now refer to RBM as MfDR to place the emphasis on development rather than organizational results).

The main objectives of good planning, monitoring and evaluation—that is, RBM— are to:

Support substantive accountability to governments, beneficiaries, donors, other partners and stakeholders, and the Management of the Plan or Project

Prompt corrective action

Ensure informed decision making

Promote risk management

Enhance organizational and individual learning

These objectives are linked together in a continuous process

1.7.   *Focus on Development Effectiveness*

Results management also means focusing on achieving development effectiveness. Meaningful and sustainable development results require more than just a generic plan of outcomes, outputs and activities. How we do development is often equally if not more important than what we do in development work. For this reason, many development agencies attempt to incorporate various themes into their planning, monitoring and evaluation processes to improve the overall effectiveness of their efforts. For example, planning, monitoring and evaluation must focus on sustainability. This conclusion was reached after years of experience with projects and programmes that had short-term impact but failed to alter the development conditions of countries or communities in any meaningful manner. Similarly, there is now a focus on gender in planning, monitoring and evaluation. Many projects and programmes often failed to achieve their objectives because there was little or no analysis of, and attention to, the differences between the roles and needs of men and women in society. Inequalities, discriminatory practices and unjust power relations between groups in society are often at the heart of development problems. The same applies to the concept of national or community ownership of development programmes. There is

greater pride and satisfaction, greater willingness to protect and maintain assets, and greater involvement in social and community affairs when people have a vested interest in something—that is, when they feel 'ownership'. Applying these principles to planning, monitoring and evaluation in a concrete manner means that these processes should be designed in such a way that they do the following:

Ensure or promote national ownership

Promote national capacity development

Promote inclusiveness, gender mainstreaming and women's empowerment

1.8.   Why Evaluate? Uses of Evaluation

Evaluation is critical to progress towards advancing human development. Through the generation of 'evidence' and objective information, evaluations enable managers to make informed decisions and plan strategically. All development project and plan success depends, in part, on the ability to carry out credible evaluations and use them to make evidenced-based decisions. The effective conduct and use of evaluation requires adequate human and financial resources, sound understanding of evaluation and most importantly, a culture of results-orientation, learning, inquiry and evidence-based decision making. Everyone in an administration and its stakeholders have to share the same vision and be open to change. When evaluations are used effectively, they support programme improvements, knowledge generation and accountability.

1.8.1.   Supporting programme improvements; did it work or not, and why?

How could it be done differently for better results? The interest is on what works, why and in what context. Decision makers, such as managers, use evaluations to make necessary improvements, adjustments to the implementation approach or strategies, and to decide on alternatives. Evaluations addressing these questions need to provide concrete information on how improvements could be made or what alternatives exist to address the necessary improvements.

1.8.2. Building knowledge for wider- use; what can we learn from the evaluation?

How can we apply this knowledge to other contexts? The main interest is in the development of knowledge for global use and for generalization to other contexts and situations. When the interest is on knowledge generation, evaluations generally apply more rigorous methodology to ensure a higher level of accuracy in the evaluation and the information being produced to allow for generalization and wider application beyond a particular context. Evaluations should not be seen as an event but as part of an exercise whereby different stakeholders are able to participate in the continuous process of generating and applying evaluative knowledge. Managers, together with government and other stakeholders, decide who participates in what part of this process (analysing findings and lessons, developing a management response to an evaluation, disseminating knowledge) and to what extent they will be involved (informed, consulted, actively involved, equal partners or key decision makers). These are strategic decisions for managers that have a direct bearing on the learning and ownership of evaluation findings. An evaluation framework that generates knowledge, promotes learning and guides action is an important means of capacity development and sustainability of results.

1.8.3. Supporting accountability; are we doing the right things?

Did the plan/project what it said it would do? The interest here is on determining the merit or worth and value of an initiative and its quality. An effective accountability framework requires credible and objective information, and evaluations can deliver such information. Evaluations help ensure that goals and initiatives are aligned with and support the Millennium Declaration, MDGs, and global, national and corporate priorities. The organization is accountable for providing evaluative evidence that links contributions to the achievement of development results in a given country and for delivering services that are based on the principles of human development. By providing such objective and independent assessments, evaluations in support the organization's accountability towards its Executive Board, donors, governments, national partners and beneficiaries. The intended use

determines the timing of an evaluation, its methodological framework, and level and nature of stakeholder participation. Therefore, the use has to be determined at the planning stage. The following Box provides a set of questions to guide and its stakeholders in assessing the potential use of evaluations. These uses are not mutually exclusive and evaluation, in general, has multiple uses. Throughout the evaluation process, the identified use has to be revisited and redefined, as necessary, in consultation with stakeholders. This inclusive process ensures the credibility and ownership of the evaluation process and products, hence resulting in its optimal use.

Box: **Assessing the use of an evaluation**

Examples of required information

Information on the relevance of intended outputs or outcomes and validity of the results framework and results map

Information about the status of an outcome and factors affecting it

Information about the effectiveness of any partnership strategy

Information about the status of project implementation

Information on the cost of an initiative relative to the observed benefits

Information about lessons learned

Who will use the information? The intended users of evaluation are those individuals or groups who have a vested interest in the evaluation results and are in a position to make decisions or take action based on the evaluation results. Users of evaluation are varied but generally fall within the following categories in the context:

The management and programme or project officers and managers, others involved in design and implementation

National government counterparts, policy makers, strategic planners

Development partners

Donors and other fund raisers

Public and beneficiaries

An organization's Executive Board and other national oversight bodies

Examples of how the information will be used

To design or validate a development strategy

1.9.   Evaluation Policy: Principles, Norms and Standards for Evaluation

In 2006 UNDP adopted an evaluation policy to strengthen the evaluation functions. The guiding principles, norms and standards as expressed in the policy and the UNEG Norms and Standards for Evaluation in the UN system guide [1] Norms for evaluation—how evaluation should be conducted in order to meet the required quality standards and its intended role—are summarized in the subsequent Box.

The remaining evaluation section of this unit aims to provide practical guidance on how these norms and principles can be applied throughout the evaluation process, focus in the training part will be on impact evaluation.

1.10. An Example: PRS (Poverty Reduction Strategy) Different Results

Indicators are typically classified into two major groups. First, final indicators measure the outcomes of poverty reduction programs (such as higher consumption per capita) and the impact on dimensions of well-being (such as reduction of consumption poverty). Second, intermediate indicators measure inputs into a program (such as a conditional cash-transfer or wage subsidy scheme) and the outputs of the program (such as roads built, unemployed men, and women hired). Target indicators can be represented as overall measurements of well-being. The so-called logic framework spells out the inputs, outputs, outcomes, and impacts in the M&E system. Impact evaluation, which is the focus of this part, spans the latter stages of the M&E framework. Viewed in this framework, monitoring covers both implementation and performance (or results-based) monitoring. Intermediate indicators typically vary more quickly than final indicators, respond more rapidly to public interventions, and can be measured more easily and in a more timely fashion. Selecting indicators for monitoring against goals and targets can be subject to resource constraints facing the project management authority. However, it is advisable to select only a few indicators that can be monitored properly rather than a large number of indicators that cannot be measured well.

1.10.1. Two propositions

We will try to compare tow documents and refer to their use of framework and indicators to analyze their different findings.

First there is a organization close to government and official aid: IFPRI, the Food Policy Research Institute Washington D.C [4]:

Module 9: Evaluation

Using 1996–97 and 2002–03 nationally representative household surveys, we examine the extent to which growth in Mozambique has been pro-poor. While all sections of society enjoyed a rapid annual increase in consumption between the sample periods, the rate of growth in consumption was slightly higher for richer households. This has led to a moderate increase in inequality at the national level, as demonstrated by the rise in the Gini coefficient from 0.40 to 0.42. However, this slight increase in inequality is not statistically significant, and its impact on poverty reduction efforts is small: the poverty headcount would have been 53.0 percent in 2002–03 if all sections of society had enjoyed the mean growth rate in consumption, compared with the 54.1 percent at which it actually stood. Interestingly, the use of the entropy class of inequality measures indicates that inequality in real consumption between provinces and regions has diminished over time, in contrast to popular claims.

Second there is a rather independent analyst to disagree with many of their theses [5], namely:

The World Bank in 2007 talked of Mozambique's 'blistering pace of economic growth'.[6] A joint donor-government study in early 2007 said 'Mozambique is generally considered an aid success story. by 'Donor cooperation strategy with Mozambique', KPMG, 2007 [7] The IMF in early 2007 said 'Mozambique is a success story in Sub-Saharan Africa. But the author claims

- People's living standards are very insecure and there is substantial movement in and out of poverty.
- Differentiation is increasing, both between the poor and better off, and within the poor (between the very poor, abjectly poor and starving).
- The claimed fall in poverty is exaggerated. People are eating more cassava, which may be a cause of malnutrition.
- Most poor people do not have the assets and access to use the present market-centred economic model, so cannot pull themselves out of poverty by their own bootstraps

1.10.2. A short comparison

There are some other sources which can be added to this example. One is a recent work of the same author: John Hanlon as before [8], which can be read under [9]. The other is a qualitative study measuring welfare and well-being, which can be read under [10].

Let us first state, that both compared documents use rather similar presentation tools:

Gini Coefficients, Lorenz curves, Theil or GE(0) coefficients

Poverty lines (more in the first case than the second)

Module 9: Evaluation

Deviation from the mean (comparing median and mean of income)

Some difference between the two is the type of questions posed:

There is a limited framework used in both cases

The IFPRI document is concerned with the national and at a lower level the provincial level.

The information sources are mainly the HH surveys 1997 and 2003

The Hanlon document relies equally on the mentioned HH surveys but includes the

Agricultural Survey (TIA) of 1996 and 2002 and focuses on the discrepancies for the rural

population. The sample design in both surveys does not allow any analysis on a lower level

than provincial

The questions posed are similar but point in a different direction. Whereas 1.) argues that the

Gini for the national level does not allow any observation of any increasing discrepancies 2.)

looks into the situation of the rural population. Here Hanlon finds that the situation has

worsened for the rural population.

But now arguing from a different viewpoint both see their viewpoint confirmed: 1.) advocates

the success of the national economic policy, reducing the number of poor by calculation of

how many have been and are now below the poverty line, Gini and Theil coefficient are used

to show increased discrepancies in urban areas but view rural areas as relatively unaffected

by any differences for different income groups. 2.) looks predominantly at the rural poor and

uses different information sources notably qualitative information about "how the situation

has changed". This leads to very different results but given the different outsets these results

are not very surprising.

Clearly none of the documents argues to be an evaluation of a political plan or program but

states their work as discussion paper, but it is intriguing to look at both documents under the

aspect of as if they were evaluations.

1.10.3. Summarizing pros and cons

Pro: Reliable sources, scientific display of data, clear relationships and diagrams

Cons: No framework, no evaluation of different objectives (Inputs, Outputs, Outcomes,

Impact), no details on clear indicators of inequality (Theil), reliance on traditional indicators,

inequality is more than consumption indicators, neglecting rural population and economy as essential for country's economical balance, baseline reference dubious (1997) because of different sample design.

Pro: Reliable sources, scientific display of data, wider use of information, creative use of discussion of inequality, use of supply factors: access to market, labour, use of economical indicators especially in rural areas, assessing rural population and economy as essential for country's economical balance, details on clear indicators of inequality (Theil),

Cons: No framework, no evaluation of different objectives (Inputs, Outputs, Outcomes, Impact), reliance on indicators tailored to demand, use of qualitative data in conjunction with quantitative data needs more justification and time-series comparison, baseline reference dubious (1997) because of different sample design.

1.11.   Methodologies to reduce bias and promote better evaluation – an Overview

Economic models can help in understanding the potential interactions—and interdependence—of a program with other existing policies and individual behaviour. Unlike reduced-form estimations, which focus on a one-way, direct relationship between a program intervention and ultimate outcomes for the targeted population, structural estimation approaches explicitly specify interrelationships between endogenous variables (such as household outcomes) and exogenous variables or factors. Structural approaches can help create a schematic for interpreting policy effects from regressions, particularly when multiple factors are at work. Ex ante evaluations, discussed in chapter 2, also build economic models that predict program impacts amid other factors. Such evaluations can help reduce costs as well by focusing policy makers' attention on areas where impacts are potentially greater. The evaluations can also provide a framework for understanding how the program or policy might operate in a different economic environment if some parameters (such as rates of return on

capital or other prices) were changed. This chapter presents case studies of different modelling approaches for predicting program effects as well as for comparing these predictions with data on outcomes after program implementation.

Thus far, the discussion has focused on ex post, data-driven evaluations of programs. Often, however, programs are implemented amid several other policy changes, which can also affect how participants respond. Creating a conceptual model of the economic environment and choices of the population in question can help in understanding the potential interactions—and interdependence—of the program with other factors. At the macroeconomic level, these factors can include other economic or social policy changes, and at the household or individual level, these factors can include different preferences or other behavioural elements. This chapter discusses structural versus reduced-form empirical approaches to estimating the causal effects of policies. It then discusses economic models in macroeconomic contexts, as well as more focused models where households face a single policy treatment, to examine how policy changes unfold within a given economic environment. Because construction of economic models is context specific, the focus is on case studies of different modelling frameworks that have been applied to various programs. When conducted before a program is implemented, economic models can help guide and streamline program design as well as draw policy makers' attention to additional, perhaps unintended, effects from the intervention.

1.11.1. Structural versus Reduced-Form Approaches

The literature discussed before centres on a single, direct relationship between a program intervention and ultimate outcomes for the targeted population. Selection bias and the problem of the unobserved counterfactual are the main identification issues addressed through different avenues, experimental or non-experimental. This one-way effect is an example of a reduced-form estimation approach. A *reduced-form approach*, for example, specifies a household or individual outcome Y_i as a function of a program T_i and other exogenous variables X_i:

$Y_i = \alpha + \beta T_i + \gamma X_i + \varepsilon_i.$ (1.10.1)

In equation 1.10.1, the program and other variables Xi are assumed to be exogenous. The *treatment-effect approach* is a special case of reduced-form estimation, in a context where Ti is appropriated to a subset of the population and Yi and Xi are also observed for separate comparison groups. The main relationship of interest is that between the policy intervention and outcome and lies in establishing the internal validity of the program's effect

1.11.2. Modelling the Effects of Policies

Because the effects of policies are very much *context* specific, coming up with a single approach for modelling policy effects in an economic framework is impossible. However, this chapter presents some basic information to help explain how these models are set up and compared with actual data. Although a number of different modelling approaches exist, here the focus is on models that examine potential price effects from policies and shocks on the utility-maximizing problem for households.

1.11.3. Assessing the Effects of Policies in a Macroeconomic Framework

Modelling the effects of macroeconomic policies such as taxes, trade liberalization, or financial regulation can be very complex, because these policies are likely to be concurrent and to have dynamic effects on household behaviour. Economic shocks such as commodity price increases or liquidity constraints stemming from the recent global financial crisis also jointly affect the implementation of these policies and household outcomes; the distributional impacts of these shocks also depend on the extent to which heterogeneity among economic agents is modelled. A number of studies (for example, [16]) have constructed general equilibrium models to examine the effects of macroeconomic policy changes and shocks on the behaviour of economic agents (households and firms) across economic sectors. These provide a useful discussion of different macroeconomic models. Again, the focus here is on a model of the price effects of policies on household utility and firms' profits.

The model equation usually is a nonlinear optimization equation. Solving these equations requires various iterations because in many cases no simple "looking for a simple peak" is possible. The mathematical approach cannot be fully extended here.

It should be added here that the pure economical model have been loosing ground constantly for economical outcomes. Social factors become more and more important and make sense to predictions and modelling. This is not a bye to a scientific picture of the world but rather to a mechanistic viewpoint, looking at humans as economic actors in a perfect economical environment (compete markets etc.) does not provide reliable predictions as the abstract reads from. [17]:

Abstract The strong interrelation between economic and social aspects suggests the necessity to include social components into macroeconomic models. Neglecting these aspects will certainly lead to misspecification and consequently unreliabilities in prediction and policy evaluation. This could be the main reason for the minor role such models are today playing in economic policy.

In this paper it is advocated to identify such missing social aspects utilizing information on individuals' happiness and the discussion on sustainability, and to include them into the models.

Such "comprehensive" models would certainly be very large. A discussion of existing modelling

1.11.4. Modelling Household Behaviour in the Case of a Single Treatment:

As discussed before, ex ante evaluations, which build economic models to predict program impacts before actual implementation, have much to offer in guiding program design as well as subsequent ex post evaluations. They can help reduce costs by focusing policy makers' attention on areas where impacts are potentially greater, as well as provide a framework for understanding how the program or policy might operate in a different economic environment if some parameters (such as rates of return on capital or other prices) were changed. Counterfactual simulations are an important part of the ex ante evaluation exercise. That is, the researcher has to construct a counterfactual sample that would represent the outcomes and other characteristics of the control group had it received the counterfactual policy.

Creating this sample requires a model to describe how the group would respond to such a policy.

1.11.5. Conclusions

This chapter has provided an overview of economic models that can shed light on mechanisms by which programs affect household choices and outcomes. Building economic models can help reduce costs by focusing policy makers' attention on areas where impacts are potentially greater. Economic models can also provide a framework ex ante for understanding how the program or policy might operate in a different economic environment if some parameters (such as rates of return on capital or other prices) were changed. However, estimation of economic models is not necessarily straightforward; careful consideration of assumptions and functional forms of equations in the model affect the estimation strategy.

Impact evaluation methods examine whether program effects can be *identified*. That is, they seek to understand whether changes in such outcomes as consumption and health can be attributed to the program itself—and not to some other cause. This chapter described major quantitative methods that are primarily used in ex post impact evaluations of programs and policies. It also discusses how distributional impacts can be measured, as well as ex ante approaches that predict the outcomes of programs and mechanisms by which programs affect targeted areas.

Randomized evaluations seek to identify a program's effect by identifying a group of subjects sharing similar observed characteristics (say, across incomes and earning opportunities) and assigning the treatment randomly to a subset of this group. The non-treated subjects then act as a comparison group to mimic counterfactual outcomes. This method avoids the problem of selection bias from unobserved characteristics. Randomized evaluations, however, may not always be feasible. In such cases, researchers then turn to so-called non-experimental

methods. The basic problem with a non-experimental design is that for the most part individuals are not randomly assigned to programs, and as a result selection bias occurs in assessing the program impact. This book discusses a number of approaches that address this problem. Propensity score matching methods, for example, attempt to reduce bias by matching treatment and control households on the basis of observable covariates. Propensity score matching methods therefore assume that selection bias is based only on observed characteristics and cannot account for unobserved heterogeneity in participation.

More can be learned if outcomes are tracked for both participants and non-participants over a time period that is deemed sufficient to capture any impacts of the intervention. A popular approach in non-experimental evaluation is the double-difference (or difference-in-difference) method, although this method can also be used in experimental approaches. This method postulates that if outcomes for both participants and non-participants are tracked over a period of time, such tracking would provide a good basis for identifying the program effect. So with double difference, the observed changes over time for non-participants provide the counterfactual for participants. Double-difference methods assume that unobserved heterogeneity is present and that it is time invariant—the treatment effect is determined by taking the difference in outcomes across treatment and control units before and after the program intervention.

An instrumental variable method identifies exogenous variation in treatment by using a third variable that affects only the treatment but not unobserved factors correlated with the outcome of interest. Instrumental variable methods relax assumptions about the time-invariant nature of unobserved heterogeneity. These approaches can be applied to cross-section or panel data, and in the latter case they allow selection bias on unobserved characteristics to vary with time. Instruments might be constructed from program design (for example, if the program of interest was randomized, or from exogenous rules in determining who was eligible for the program), as well as from other exogenous shocks that are not correlated with the outcomes of interest.

Module 9: Evaluation

Regression discontinuity and pipeline methods are extensions of instrumental variable and experimental methods that exploit exogenous program rules (such as eligibility requirements) to compare participants and non-participants in a close neighbourhood around the rule's cut-off point. Pipeline methods, in particular, construct a comparison group from subjects who are eligible for the program but have not yet received it.

Although experimental methods are, in theory, the ideal approach for impact evaluation, non-experimental methods are frequently used in practice either because program administrators are not too keen to randomly exclude certain parts of the population from an intervention or because a randomized approach is out of context for a rapid-action project with no time to conduct an experiment. Even with an experimental design, the quality of impact analysis depends ultimately on how it is designed and implemented. Often the problems of compliance, spillovers, and unobserved sample bias hamper clean identification of program effects from randomization. However, non-experimental methods such as propensity score matching, double difference, and use of instrumental variables have their own strengths and weaknesses and hence are potentially subject to bias for various reasons including faulty design of the evaluation framework.

This chapter also covers methods of examining the distributional impacts of programs, as well as modelling approaches that can highlight mechanisms (such as intermediate market forces) by which programs have an impact. Well-being can be assessed at different levels, for example, among individuals or households, as well as for geographic areas such as villages, provinces, or even entire countries. Impacts can also be differentiated more finely by gender, percentiles of income, or other socio-economic or demographic characteristics. Factoring in nuances of program effects, either across the distribution of income or through models of market interactions can help in understanding the mechanisms of the program's effects as well as in reducing costs by focusing policy makers' attention on areas where impacts are potentially greater.

In reality, no single assignment or evaluation method may be perfect, and verifying the findings with alternative methods is better. Different ex ante and ex post evaluation methods

can be combined, as can quantitative and qualitative approaches. The main lesson from the practice of impact evaluation is that an application of particular methods to evaluate a program depends critically on understanding the design and implementation of an intervention, the goals and mechanisms by which program objectives can be achieved, and the detailed characteristics of targeted and non-targeted areas. By conducting good impact assessment over the course of the program and by beginning early in the design and implementation of the project, one can also judge whether certain aspects of the program can be changed to make it better. More information and comprehensive examples in [2]

The rather technical description of the evaluation will continue with recommendations about how to address the stakeholders

1.12.   Addressing stakeholders and the public

Different groups of stakeholders will need and require different types of information:

Politics, decision makers: Will need timely information about performance on project goals, and pro

Let us refer to other modules an relate the possible (monitoring) and evaluation to the different layers of a M&E framework and the inventories for the different levels

Stakeholders / Levels	Government, Politics		Project/ Plan implementers and Institutions	Scientific Community	Public
the entry conditions as for instance the Government and Donor Readiness	Wrap-up notice No details	System (MIS) Management Information	Technical evaluation (light)	Technical evaluation (see above)	Information about the entry conditions in concise journalistic form

Module 9: Evaluation

Stakeholders ╲ Levels	Government, Politics		Project/ Plan implementers and Institutions	Scientific Community	Public
the immediate effects/activities	None		Statistics / Graphs	Statistical Analysis	Journalistic Samples
the outputs	Selective Journalistic coverage		Methodological Tools output	Output Analysis (see above)	Selective Journalistic coverage
the outcomes	Simplified Outcome Analysis		Methodological Tools outcome (ex-post & ex-ante)	Outcome Analysis (see above)	Comprehensive Journalistic coverage
the impact	Simplified Impact Analysis		Technical outlook and projection (light)	Impact Analysis. Technical outlook and projection	Journalistic Reporting

This is obviously a simplistic and schematic display of how and stakeholders should receive as information. Project Managers, Line managers, Administrators are sub summarized as implementers in above scheme and are likewise stakeholders on the giving and the receiving end.

We will continue with more precise and illustrating examples mostly from [3] but first some general considerations:

1.12.1. Conceptualizing utilization and influence

When assessing the use and utility of an evaluation, it is helpful to consider two components: we term them "utilization" and "influence."

Module 9: Evaluation

Utilization: How were the evaluation findings (and even the process) used - by whom and for what purpose? The first uses that generally come to mind are those related to impact evaluation as an assessment tool. For example, one may conduct an evaluation in order to:

monitor project implementation,

measure the benefits of an existing program and check for unanticipated side effects,

assess the distribution of participation and benefits across different segments of the target population,

make informed changes and improvements to an ongoing project,

test options for the design of a project that will be implemented in the future, and

compare the cost-effectiveness or benefit/cost ratio of alternative programs for budget and planning purposes.

In practice, however, impact evaluations are also very commonly used as a political tools and the better they are communicated the more this will be the case.

They are frequently employed to:

provide support for decisions that agencies have already decided upon or would like to make

mobilize political support for high profile or controversial programs,

provide independent support (the international prestige and perceived independence of the evaluator is often important) for terminating a politically sensitive program, and

provide political or managerial accountability.

In fact, in the end it is likely to be the potential political benefit or detriment that causes decision makers to embrace or avoid evaluations. As a result, those who would like to promote impact evaluation as an assessment and learning tool will have to be fully aware of the given political context and navigate strategically.

Influence: In assessing the influence of an impact evaluation, there are a number of aspects one might consider:

Module 9: Evaluation

What causes or facilitates an impact evaluation's influence

Where can the evaluation's influence be seen?

How much influence did the evaluation have on the decisions and actions of managers,
planners and policymakers?

1.12.2. Reviewing the evidence

How were the evaluations utilized and what kinds of influence did they have ?

In the following there is a summary of the utilization and influence that were observed in the
12 case studies.

In most cases the information is based on the perceptions and experience of the evaluators
themselves, although in one case a representative of the government client agency was also
present. Not surprisingly you will find one successful evaluation example and two rather less
successful but not untypical examples.

The reason to illustrate the less efficient and successful evaluations is NOT to criticize their
authors. This would be arrogant and neglecting circumstances and possible limitations. It is
rather to show the shortcomings which may render the effort of an costly and well intended
evaluation useless and of little help for the intended purposes of the same evaluation.

The types of use and influence seen in the presented cases can be broadly grouped into
three categories:

Project implementation and administration,

Evaluations were often used for the design of future programs. They provided specific
operational guidance or general guidance for the strategic focus. They often helped identify
logistical and administrative problems that had been overlooked

Political support, and the process of evaluation

Module 9: Evaluation

Evaluations are often used to justify continued funding for a program, or to ensure political support for a new or expanded program. In several cases an evaluation was used to justify a new program, even when in fact the evaluation findings did not support this new program

Culture of and capacity for impact evaluation

Evaluations that are favourably received by clients often lead to increased interest in further evaluations. Well designed and implemented evaluations have helped legitimize evaluation as a useful planning or policymaking tool. Initially many clients or local districts were either sceptical about an evaluation's utility or were afraid that the findings would be too negative or critical. In several cases attitudes became more positive and utilization increased as the evaluations progressed. Not surprisingly, it was much easier to gain acceptance for the evaluation process and findings when the findings were mainly positive. Well received evaluations often lead to follow-up evaluations to assess more specific issues that had been identified.

There were, however, examples, where initial negative findings created reluctance to accept or use an evaluation, but where attitudes gradually became more favourable. The health insurance evaluation in China (see examples) was very poorly received in the beginning because it showed negative results on the primary objective of reducing out-of-pocket health care expenditures (though positive results for a secondary objective of increasing use of health care services). In the end, though, authorities accepted the results and were able to use them to make some reforms (especially increased funding), and the process seemed to have increased general acceptance of impact evaluation as a tool.

.

1.13.   Maintaining awareness (The influence and utilization of impact evaluations)

Created demand for further and more	Demand for follow-up studies Demand for new evaluations Demand for more rigorous evaluations Demand to develop evaluation systems for particular sectors Demand for methodologies to evaluate pilot projects Appreciation of the value of independent, external evaluations

Strengthening impact evaluation methodology	Development of more rigorous evaluations Strengthened Management Information System (MIS) and data quality Demonstrated the value of contextual analysis Introducing evaluation capacity development Institutionalization of evaluation

Strengthening program design and implementation	Strengthening the design of future projects Assessing the cost-effectiveness of alternative interventions Implementation and evaluation of pilot projects before launching major

Providing evidence to support programs and	Providing evidence to respond to critics Justifying continuation of program Justifying continued inclusion of particular component Raised profile of a program/intervention

Providing evidence to challenge a program	Used by donors to challenge program they do not agree with. This involves wider group of stakeholders

2. Using the Results

2.1.   Factors affecting evaluation utilization and influence

The following is a synthesis of the broad range of factors identified in the presentations as potentially affecting evaluation utilization.

2.1.1.   Timing and focus on priority stakeholder issues

The evaluation must be timely and focus on priority issues for key stakeholders. This ensures there is a receptive audience. Timing often presents a trade-off: on the one hand, designing an evaluation to provide fast results relevant for the project at hand, in time to make changes in project design and while the project still has the attention of policymakers. On the other hand, evaluations that take longer to complete may be of higher quality and can look for longer term effects on the design of future projects and policies.

The evaluator must be opportunistic, taking advantage of funding opportunities, or the interest of key stakeholders. Several countries that have progressed toward the institutionalization of evaluation at the national or sector level began with opportunistic selection of their first impact evaluations.

The evaluator should always be on the look-out for "quick-wins" – evaluations that can be conducted quickly and economically and that provide information on an issue of immediate concern. Showing the practical utility impact evaluations can build up confidence and interest before moving on to broader and more complex evaluations. Also, there is value in firsts. Pioneer studies may not only be useful for showing the impact of the intervention, but in a broader context they may also change expectations about what can and should be evaluated or advance the methods that can be used. Again, even less-than-ideal evaluations that are first or early in their context may contribute by building interest in and capacity for impact evaluation.

A series of sequential evaluations gradually builds interest, ownership and utilization.

2.1.2. Effective dissemination

Rapid, broad and well targeted dissemination are important determinants of utilization. One reason that many sound and potentially useful evaluations are never used is that very few people have ever seen them.

Making data available to the academic community is also an important way of broadening interest and support for evaluations and also of legitimizing the methodologies.

Providing rapid feedback to government on issues enhances utilization.

Continuous and targeted communication builds interest and confidence and also ensures "no surprises" when the final report and recommendations are submitted. This also allows controversial or sensitive findings to be gradually introduced. Trust and open lines of communication are important confidence builders.

Where there is existing demand for a particular evaluation, the results may be partially disseminate themselves and may be more likely to be used.- How to raise demand that is the question

2.2. Feedbacks, when and where

2.2.1. Clear and well communicated messages

Clarity and comprehensibility increase use. It helps when the evaluation results point to clear policy implications. This may also apply to the comprehension of methods. While stakeholders may be willing to "trust the experts" if an evaluation offers results that support what they want to hear, there may be a reasonable tendency to distrust results – and particularly methods – that they don't understand.

2.2.2. Active engagement with national counterparts

The active involvement of national agencies in identifying the need for an evaluation, commissioning it, and deciding which international consultants to use is central to utilization.

Close cooperation with national counterpart agencies proves critical in several ways. It gives ownership of the evaluation to stakeholders and helps ensure the evaluation focuses on important issues. It often increases quality by taking advantage of local knowledge and in several cases reduces costs (an important factor in gaining support) by combining with other ongoing studies. This cooperation can enable evaluators to modify the initial evaluation design to reflect concerns of clients – for example, changing a politically sensitive randomized design to a strong quasi-experimental design.

Involving a wide range of stakeholders is also an important determinant of utilization. This can be achieved through consultative planning mechanisms, dissemination and ensuring that local as well as national level agencies are consulted.

In many contexts, the involvement of the national statistical agency increases the government's trust – the results and the process have been better accepted when overseen and presented by the statistics agency.

2.2.3.   Demonstrating the value of evaluation as a political and policymaking tool

When evaluation is seen as a useful political tool, this greatly enhances utilization. For example, managers or policymakers often welcome specific evidence to respond to critics, support for continued funding or program expansion. Evaluation can also be seen as a way to provide more objective criticism of an unpopular program.

Once the potential uses of planning tools such as cost-effectiveness analysis are understood, this increases the demand for, and use of, evaluations. Evaluations can also demonstrate the practical value of good monitoring data, and increased attention to monitoring in turn generates demand for further evaluations. When evaluations show planners better ways to achieve development objectives, such as ensuring services reach the poor, this increases utilization and influence.

Increasing concerns about corruption or poor service delivery have also been an important factor in government decisions to commission evaluations. In some cases, a new administration wishes to demonstrate its transparency and accountability or to use the evaluation to point out weaknesses in how previous administrations had managed projects.

411

Evaluations that focus on local contextual issues (i.e. that are directly relevant to the work of districts and local agencies) are much more likely to be used.

2.3. Assuring Quality in the evaluation

2.3.1. The methodological quality of the evaluation and credibility of the international evaluators

High quality of an evaluation is likely to increase its usefulness and influence. Quality improves the robustness of the findings and their policy implications and may assist in dissemination (especially in terms of publication). However, an impact evaluation of a compromised quality may still be useful if it can provide timely and relevant insight or if it ventures into new territory: new techniques, less-evaluated subject matter, or in a context where relevant stakeholders have less experience with impact evaluations.

The credibility of international evaluators, particularly when they are seen as not tied to funding agencies, can help legitimize high profile evaluations and enhance their utilization. In some cases the use of what is considered "state of the art" evaluation methods, such as randomized control trials, can raise the profile of evaluation (and the agencies that use it) and increase utilization.

New and innovative evaluations often attract more interest and support than the repetition of routine evaluations. On the other hand, while studies on the "frontier" may be more novel or attract more attention, subsequent related studies may be useful in confirming controversial findings and building a body of knowledge that is more accepted than a single study, especially a single study with unpopular findings.

Evaluation methods, in addition to being methodologically sound, must also be **understood and accepted** by clients. Different stakeholders may have different methodological preferences and capacities (see above).

2.3.2. Positive and non-threatening findings

Positive evaluations, or those that support the views of key stakeholders, have an increased likelihood of being used. While this is not surprising, one of the reasons is that many agencies were either fearful of the negative consequences of evaluation or (to be honest) considered evaluation as a waste of time (particularly the time of busy managers) or money. Once stakeholders have appreciated that evaluations were not threatening and were actually producing useful findings, agencies have become more willing to request and use evaluations and gradually to accept negative findings – or even to solicit evaluations to look at areas where programs were not going well.

There is always demand for results that confirm what people want to hear. There may be some benefit in taking advantage of opportunities to present good results, especially if it helps the process of getting stakeholders to understand and appreciate the role of impact evaluation. Sometimes, though, demand can be built despite less-positive results – by special efforts to target the relevant stakeholders. Concerns over potential negative results, bad publicity, or improper handling of the results may reduce demand; sensitivity, trust-building, and creative arrangements may help overcome these fears.

2.3.3. Evaluation as a fig-leaf for political continuation

In not few cases, the evaluation of a program's or development plan's success is essential for a political continuation. Poverty reduction is a very good example. Donor will only be willing to continue contributing to national PRS if success has been "proven". This makes an independent evaluation almost impossible. The only solution is a mixture of increasing ownership and increasing independence of impact evaluations.

In this context it must be avoided that institutes, agencies continue to evaluate on a perpetual basis. However, this is very difficult to avoid, because the established technical liaisons. Any major development measure does not only need a monitoring system to e established at the project definition stage but also an independent evaluation strategy to be commissioned for the impact evaluation

2.4.   Evaluation: a cyclical model

(see initial graph)

2.4.1.   Evaluation capacity development

Evaluation capacity, especially at a local level, is an important factor in the quality of any impact evaluation that also affects the ability of stakeholders to demand, understand, trust, and utilize the results.

Capacity building is an iterative process and may improve both demand and quality.

2.4.2.   Pursuing easy wins alongside harder challenges

The most effective strategy for developing a strong culture of evaluation may be two-pronged: opportunism where there are "easy wins" – willing partners, high capacity, good data, good results, etc., since these may require less effort and fewer resources and may generate familiarity with the process; and at the same time "chipping away" systematically at the harder problems where there is less capacity or less tradition of evaluation.

2.5.   3 Scenarios of Impacts (Examples of Development Programs and Evaluations)

All scenarios are from [3], explained in detailed there ad discussed here for their exemplary standing in the next Module

Program	Evaluation questions	Main findings
EDUCATION PROGRAMS		
Uganda: Universal Primary Education (UPE) **Goals:** Test the effectiveness of improved management	■ Trends in attendance and learning since 2000 ■ Determinants of trends ■ Size and cost-effectiveness of each intervention ■ Use of MIS for evaluation	■ Progress in access to education ■ Effectiveness of investments in teachers, classrooms, books and other facilities ■ School management important ■ Investments more effective if combined with improved management ■ Quality of primary education remains poor and absenteeism and drop-outs high

Module 9: Evaluation

Program	Evaluation questions	Main findings
CONDITIONAL CASH TRANSFERS [CCT] AND POVERTY REDUCTION PROGRAMS		
Progresa/ Oportunidades: Mexico. Conditional cash transfers promoting children's health, nutrition and education. **Goals:** Short-term poverty reduction through cash transfers. Long-term investment in human capital development through increasing access to health and education	■ Are CCTs cost-effective in increasing access of poor children to health and education? *Effectiveness of key program components:* ■ Direct monetary transfers versus in-kind grants ■ Targeting the extremely poor versus all families ■ New, standard targeting procedures versus existing program client lists ■ Transfers to households versus to communities ■ Non-discretionary rules for whole country versus flexibility for local authorities ■ Directing benefits directly to women versus to household head ■ Program impacts on fertility ■ Criteria for defining size of transfer ■ Merits of family co-responsibility and certification	■ Poverty targeting worked well [10] ■ PROGRESA reduces by 10% people living below poverty line ■ Positive impact on school enrolment for boys and girls ■ Children entering school earlier, less grade repetition and better grade progression ■ Younger children have become more robust against illness ■ Women's role in household decision-making increases ■ Estimated cost-benefit ratio of 27%
HEALTH		
Kenya: Bed net distribution experiment: Free vs. Cost-Recovery **Goals:** *Increased distribution and use of insecticide-treated nets*	■ Is free distribution or cost-recovery more effective for increasing distribution and use of nets? ■ How price elastic is demand?	■ Cost recovery did not increase distribution or use ■ Cost recovery appears to reduce demand

---

[10] The PROGRESA findings were not reported in the conference but were taken from IFPRI (2002) PROGRESA: Breaking the Cycle of Poverty

3. Lesson for Better Evaluation

The examples and reports are mainly adapted from [3], further detailed information can and should be solicited there. This is not to show that these evaluations are of exemplary quality, quite the contrary and remarks are added when necessary. But this overview will give a good room for evaluation of evaluation. Of course, this is meant to be ironic because no such thing is intended but to learn about and prepare for better Monitoring for better Evaluation is the goal.

3.1.   Benchmarks for evaluations

3.1.1.   Impact Evaluation of Primary Education in Uganda

The program

The purpose of the evaluation was to assess the effectiveness of a number of interventions introduced into the primary education system between 2000 - 2006 and contributing to the national goal of Universal Primary Education. The interventions included: management improvements, infrastructure, teaching materials and increased number and quality of teachers. These interventions form part of the national "full coverage" education services but were also tested in more depth in the Masindi District Education Development Project.

The evaluation

The central evaluation questions were: How have school attendance and learning achievement developed since 2000? What were the main determinants of these developments? Which interventions have the largest and most cost-effective impact on educational outputs? How effectively has the Management Information System been used for purposes of evaluation?

Module 9: Evaluation

The evaluation was conducted at two levels: nation-wide and in the Masindi District. The evaluation was based on a program theory intervention model that identified four sets of interventions (school management, infrastructure, teaching materials and teachers that would enhance school performance through improving access and learning achievement; and in turn produce a set of welfare outcomes. The outcomes would be affected by local contextual factors that could affect results in each district. Given the countrywide coverage of the education programs, the many different donors and agencies involved and the large number of contextual factors in each region, it was difficult to define a counterfactual. So a number of different approaches were used: combining different data bases to increase the range of variables included in the analysis, using triangulation[11] to obtain independent estimates of key indicators, and using natural restrictions (e.g., remote rural areas where well educated parents do not have a choice of selecting schools with smaller class sizes); and propensity score matching to create ex-post comparator groups comparable with the intervention groups. In Masindi, a quasi-experimental design was used where schools receiving the project interventions were compared both with a comparator group from outside the district and with schools in the district that did not participate in the project.

3.1.2.   The evaluation findings

The main findings of the evaluation were the following:

Uganda has made enormous progress in improving access to primary education.

The analysis confirmed the effectiveness of investments in teachers, classrooms, books and other school facilities. It also confirmed that high pupil-teacher ratios and high pupil-classroom ratios have a negative effect on learning achievements.

There are also significant effects from teacher education and training.

---

[11] In the social sciences, triangulation is often used to indicate that more than two methods is used in a study with a view to double (or triple) checking results. This is also (and better) called "cross examination". Source: wikipedia.org

Head teacher qualification is also important.

Investments in teachers, classrooms and books are more effective when combined with improvements school and district management. Privately funded schools, which are generally better managed, outperform government schools by 40%.

The quality of primary education remains poor and absenteeism and dropout pose serious threats to the efficiency and effectiveness of primary education. (here reasons for this would be a far-leading indicator)

The in-depth evaluation of the Masindi District Project found that educational performance in project schools were 50-60 per cent better than the comparator group from surrounding districts, and 35% better than other schools in Masindi.

3.1.3. Impact of the evaluation

At the national level, this was the first time the Ministry of Education could respond to Parliament providing concrete evidence of the impacts and cost-effectiveness of the education programs, and refuting criticisms that the money would have been better spent on other social programs. In particular, the evaluation showed that improved management could have a greater impact on education outcomes than simply building more classrooms and hiring more teachers. *By providing an objective basis for engagement with policy makers and implementers, the evaluation encouraged the involvement of a wider range of stakeholders in education sector activities. (How was this done, is not clear. Indeed, the goal of this program is strangely vague, are we trying to improve MIS or evaluation methods or education in Uganda)*

The evaluation has also improved the quality and effectiveness of the Education Management Information System (EMIS). Demonstrating how the information can be used in an evaluation has encouraged central agencies and local authorities to improve the quality of the data they collect. (Why also, this was the primary gaol of the programme). The evaluation

also demonstrated the importance of contextual analysis to complement and go beyond the statistical data to understand the particular characteristics of each region and how these affect educational performance.

In the Netherlands, the report was published and sent to the parliament. The results of the report were used in the Netherlands in an extensive evaluation of (Dutch) Africa policy in 2008. One of the workshops of the conference confirmed the importance of management in schools. Also in the Netherlands, there has been a discussion of the low level of achievements in primary schools in Uganda. (Much more important would have been to follow the discussion in Uganda's parliament)

These findings coming out of the impact evaluation have grounded this broader "quality of primary education" discussion, linking demands to improve pupil and teacher attendance and the reduction of absenteeism to an improvement of the management in schools. In both countries, the evaluation has contributed to an interest on impact evaluation as a management tool. In Uganda, the evaluation contributed to the mentioned initiative to enhance the quality of primary education with impact evaluation as one strategy for evidence based policy formulation and decision-making.

3.1.4. Comment on Impact Evaluation of Primary Education in Uganda

The final outcome for Uganda as formulated in [3] was the implementation of several new evaluations, like an evaluation which analysed the impact of primary education on the future of boys and girls through further education and employment opportunities. If this is the only goal of an evaluation, than certainly this is not enough, political implication, demand for further other improved UPE and a discussion like in the Netherlands parliament would be necessary. Certainly the report does not say anything about this and falls short concerned to may criteria of sustainable, government owned program.

3.2. Participation of stakeholders, recommendations and limits

Again an example of evaluation is presented from [3]. Comments are added likewise

3.2.1. Impact Evaluation in the PROGRESA/ Oportunidades Program of Mexico

The program

PROGRESA, now renamed Oportunidades, is a conditional cash transfer (CCT) program that provides cash directly to low-income families on the condition that children attend school regularly and family members visit health centres. PROGRESA was one of the first CCT programs in Latin America and was influential in the design of later programs in other countries. The program had two main objectives: to produce short-run effects on poverty through cash transfers, and to contribute to long-run poverty alleviation through investment in human capital (i.e. education, health and nutrition). The focus is on children because early interventions have much higher returns over the life-cycle. Payments were made to the mother to increase the likelihood that children would benefit. The program included a number of innovative features, several of which were considered quite controversial at the time, and all of which were assessed by the evaluations. Some of the measures included:

direct monetary transfers instead of providing vouchers or food in-kind, or improving supply side services;

the programs targeted the extremely/structurally poor rather than all families;

PROGRESA developed a single national roster of beneficiaries rather than working from existing lists;

transfers were given directly to households rather than communities;

uniform, non-discretionary rules were introduced for the whole country; and

there was a requirement of family co-responsibility and certification.

Module 9: Evaluation

The evaluation

The program began one year before the 1999 Presidential elections and there was pressure from the ruling party (PRI) to ensure that the findings of the evaluation would be available prior to the election. When Vicente Fox was elected in 2000, the new administration continued to support a rigorous, independent evaluation to provide objective evidence that their programs were more effective and transparent than those of the PRI regime that had been in power for the previous 80 years. The rigorous and expensive evaluation systems were justified on three grounds:

Economic: to improve the design and effectiveness of the programs and to compare impacts and cost effectiveness of different programs;

Social: Increasing transparency and accountability, and

Political: the evaluations increased the credibility of the programs,

and this, combined with increased transparency and accountability, helped break with past practices, such as political influence in beneficiary selection.

The evaluation design

The program was implemented in phases, and for each phase beneficiary communities were selected randomly, with non-selected communities providing a non-biased comparator group. Randomization was politically acceptable because communities not selected in one phase were likely to be included in the next phase. Also the government was strongly committed to the use of rigorous, state-of-the art evaluation design to ensure credibility of the findings. 24,077 households were interviewed in 320 treatment and 186 control communities. Families were interviewed at the start of the program and at several points during implementation, avoiding problems of linear extrapolation when only one post-test measurement is made. (One can imagine, that this was quite a costly evaluation)

3.2.2. Main findings of the evaluation

The following are some of the findings highlighted in a 2002 IFPRI report on the first post-test evaluation[12]:

Poverty targeting worked well.

PROGRESA reduced by 10% (the number of) people living below the poverty line.

Positive impact on school enrolment for boys and girls.

Children entered school earlier, with less grade repetition and better grade progression.

Younger children have become more robust against illness.

Women's role in household decision-making increases.

The program had an estimated cost-benefit ratio of 27%.

3.2.3. Evaluation utilization and influence

According to the presentation, the evaluation had the following kinds of influence:

Continuation of the program under a new administration. The independence, credibility and positive outcomes of the early stages of the evaluation significantly contributed to the program's continuation under the new administration.

Improved operational performance. The early operations reports identified implementation issues, such as delivery of food supplements and intra-household conflicts, and issues with targeting rules that were addressed as the program evolved.

---

[12] IFPRI (2002) PROGRESA: Breaking the Cycle of Poverty.

Contributed to program expansion to urban areas. A youth job creation program created income generating opportunities for poor households through preferential access to micro-credit, housing improvements, adult education and social/health insurance.

Contributed to the development of a more systematic policy evaluation approach in Mexico. This move was formalized by the creation of CONEVAL (Council for Program Evaluation) in 2006.

Enhanced policy evaluation internationally. The evaluation findings were available on the internet and were used widely by academic researchers. The design and findings were able to withstand critical scrutiny, greatly enhancing the credibility and influence of the findings.

Contributed to the initiation of CCT programs in many other countries. Similar programs, most of which have been influenced to some extent by PROGRESA, have been started in at least 10 other countries.

3.2.4. Comment on Impact Evaluation of PROGRESA/ Oportunidades Program in Mexico

The final outcome for Mexiko as formulated in [3] has confirmed to reach several essential goals notably the poverty. Another essential aspect is the continuation and amplification of evaluations (and certainly monitoring) efforts. This is and can be seen a very political decision. Without any political implication of the goals and their evaluation, there would be no reason for any further follow-up or far leading political decisions. There seem to be significant results of the results of Impact Evaluation as directly correlated to the governance issues (responsibility to the public, proving the success of political decisions) but also related to technical aspects (state of the art evaluation, inclusion of the scientific community, international relationships and expansion). One notices the substantial financial effort contributed to this evaluation, quality, effectiveness and costs of the evaluation are certainly related.

3.2.5.   Evaluation of insecticide-treated nets in Kenya

The program and evaluation

This study explored the relative benefits of free distribution and cost recovery practices for maximizing coverage and usage of health products – specifically anti-malarial insecticide-treated nets. In particular, there was interest in the competing effects of higher prices reducing willingness or ability to pay and higher prices increasing the perceived value of the product, potentially reducing resource wastage. The experiment randomized the price of nets offered in 20 prenatal clinics from zero to 40 shillings ($.60), subsidizing the price by 100 to 90 percent, and then compared the uptake and usage of the nets at various prices. The primary evaluation interests were the effects of the various prices on demand for acquiring the nets and on usage a few months after acquiring a net. Uptake and usage rates were multiplied together to measure "effective coverage". The evaluation did not explore any potential loss of quality or service that might occur in association with eliminating cost-sharing outside of the experimental framework.

The evaluation findings

The evaluation found that demand drops very quickly as prices increase, and the highest price offered in the experiment is still lower than the common cost-sharing price. On the other hand, usage didn't vary much across price paid. Combining uptake and usage, free distribution led to a 63 percent coverage rate compared to 14 percent for the highest price group. Women who paid for the nets were not found to have worse health at the time of the prenatal visit, suggesting that ability to pay may be more of a limiting factor than need or willingness to pay and thus that full subsidies may be most effective in maximizing the effective coverage rate. However, it was noted that in Kenya, where the experiment was

carried out, extensive efforts have resulted in most of the population being familiar with the benefits of nets.

3.2.6. Evaluation utilization and influence

At the time of the presentation, dissemination was still in the early phases. The results were first presented to the Ministry of Health. The news was well received, as they already preferred free distribution, but they noted that they would have to work to find funding and to convince donors and NGOs in the area. To some degree, the evaluation has disseminated itself because of high existing demand for the evidence. An example was DFID contacting the authors before the paper was finished, to help decide whether to give or sell nets in Somalia, so it had immediate impact on other projects, as well as on the organization's official views. There seems to have been a mixed response in private foundations. In particular, there were rumours of methodological critiques from a major net-distributing organization. (Which one would have been interesting?) From the local branch of the same organization, however, feedback was received to say that they were really pleased with the evaluation and its results. They were changing their model to dispense nets for free, and the evaluation would help them defend their choice. Since then, the local branch has helped disseminate the study. This presentation noted some of the broader influence of impact evaluations may be harder to track, such as when evaluations contribute to a larger body of evidence. An individual evaluation may not be entirely conclusive, but conclusions drawn from an accumulation of evidence may be more difficult to refute.

3.2.7. Lessons learned

This experience showed that when an evaluation addresses an existing demand for evidence, the results may partially disseminate themselves, as interested audiences seek out

the information. Also, it seems that people tend to trust or distrust evidence based on what they already believe, looking for results that confirm what they believe and looking for ways to discredit contrary information. Perhaps one reason is that it is difficult to distinguish between good and bad evidence. Currently, there is much ongoing work to provide training in measurement and evaluation for donors and policymakers: when individuals have a greater understanding of impact evaluation, they may be better able to recognize differing qualities of evidence they come into contact with, allowing individual evaluations to have greater impact. Similarly, people may not trust evidence (especially evidence contrary to their beliefs) that comes from methods they do not understand, so training in or exposure to impact evaluation as well as the use of easy-to-understand methods may make evaluation results more convincing. In the end, an individual evaluation may not be entirely conclusive, but conclusions drawn from an accumulation of evidence may be more difficult to refute.

3.2.8. Comment on Evaluation of insecticide-treated nets in Kenya

The findings in the report are understandably reluctant to criticize, but being more outspoken as independent commentator: This evaluation could as well have been skipped.

As results there are

No significant results: Selling is just as good as giving nets away

No impact analysis: What did it mean for future or current programs

There were methodological shortcomings

3.3. Comparing evaluation reports and learning from this

The reason to discuss the evaluations is not to criticize but to learn and to implement better monitoring for better evaluation.

3.3.1. How is impact evaluations used?

Impact evaluations can be used as an *assessment tool* to help strengthen project and program design by providing a more systematic, rigorous, and quantifiable assessment of how a project has performed, what it has achieved (compared to its intended objectives), who has and has not benefited, and how the costs of producing the benefits compare with alternative ways of using the resources. Impact evaluations are also used as a *political tool* to provide support for decisions that agencies have already decided upon or would like to make, to mobilize political support for high profile or controversial programs and to provide political or managerial accountability. This latter function has been important in countries where new administrations were seeking to introduce transparency into the design and implementation of high profile, politically attractive programs. Impact evaluations can also provide independent corroboration and political cover for terminating politically sensitive programs – in which case the international prestige and independence of the evaluator was found to be important. In fact, in the end it is likely to be the potential political benefit or detriment that causes decision makers to embrace or avoid evaluations, and those who would like to promote impact evaluation as an assessment and learning tool will have to be fully aware of the given political context and navigate strategically.

3.3.2.   What kinds of influence can impact evaluations have?

The twelve impact evaluations discussed in [3] were utilized and had influence in three broad areas: project implementation and administration; providing political support for or against a program; and promoting a culture of evaluation and strengthening national capacity to commission, implement, and use evaluations. It is not only the findings of an impact evaluation that can have an impact. The decision to conduct an evaluation, the choice of methodology, and how the findings are disseminated and used can all have important consequences – some anticipated, others not; some desired and others not. For example, the decision to conduct an evaluation using a randomized control trial can influence who benefits from the program, how different treatments and implementation strategies are

prioritized, what is measured, and the criteria used to decide if the program had achieved its objectives.[13]

The influence of evaluations can be seen in administrative realms such as program design and scope or the political realm in the form of popular support for a program or its associated politicians. Understanding the role of impact evaluation is also a process that evolves as managers, policymakers and other stakeholders become more familiar with how evaluations are formulated, implemented and used. For high profile programs, the influence of the evaluation may also be seen in how the debate on the program is framed in the mass media.

3.3.3. Guidelines for strengthening evaluation utilization and influence

The following is a synthesis of the broad range of factors identified in the presentations as potentially affecting evaluation utilization.

Timing and focus on priority stakeholder issues:

The evaluation must be timely and focus on priority issues for key stakeholders. Timing often presents a trade-off: on the one hand, designing an evaluation to provide fast results relevant for the project at hand, in time to make changes in project design and while the project still has the attention of policymakers. On the other hand, evaluations that take longer to

---

[13] A frequently cited example from the US was the decision to assess the performance of schools under the

No Child Left Behind program in terms of academic performance measured through end-of-year tests.

This meant that many schools were forced to modify their curricula to allow more time to coach children in

how to take the tests, often resulting in reduced time for physical education, arts, and music.

complete may be of higher quality and can look for longer term effects on the design of future projects and policies.

Cooperation with the "clients" of an evaluation cannot begin too early. In this case, involving the government in the choice of survey design helped to ensure there was comfort with the evaluation methods and eventually the results – increasing utilization.

The evaluator must be opportunistic, taking advantage of funding opportunities, or the interest of key stakeholders. The evaluators must work closely with national counterparts and be responsive to political concerns. Several countries that have progressed toward the institutionalization of evaluation at the national or sector level began with opportunistic selection of their first impact evaluations[14].

The evaluator should always be on the look-out for "quick-wins" – evaluations that can be conducted quickly and economically and that provide information on issues of immediate concern. Showing the practical utility of impact evaluations can build up confidence and interest before moving on to broader and more complex evaluations.

There is value in firsts. Pioneer studies may not only show the impact of the intervention, but in a broader context they may also change expectations about what can and should be evaluated or advance the methods that can be used. Even less than- ideal evaluations that are the first or early in their context can build interest and capacity for impact evaluation.

---

[14] See IEG (2008) Institutionalizing Impact Evaluation within the Framework of a Monitoring and

Evaluation System. The Education for All evaluations in Uganda were cited as an example of institutionalization at the sector level and the SINERGIA evaluation program under the Planning

Department in Colombia is an example of institutionalization of a national impact evaluation system . The

report is available at: [16]

A series of sequential evaluations gradually builds interest, ownership and utilization. For impact evaluations, *there is often a trade-off between speed and quality*. The impact of higher quality evaluations may not be seen in the actual intervention being evaluated, but the benefits may extend to other projects and evaluations by pushing the frontier of what can be evaluated and how and by setting new expectations for evaluation quality.

Starting the evaluation data collection late in the project cycle may reduce data quality and the insights that could be gained. However, with careful management and particular attention to communication and dissemination, even a less-than-ideal evaluation can prove to be very useful, especially if it is timely and addresses urgent questions.

Also, there may be special challenges that can suppress demand in sectors and regions that are less commonly evaluated. Where there is little habit or "culture of evaluation", there may be less funding and less pressure to evaluate, and perhaps higher resistance to accountability. It may require special efforts to begin to build a culture of evaluation.

Clear and well communicated messages

Clarity and comprehensibility increase use. It helps when the evaluation results point to clear policy implications. This may also apply to the comprehension of methods. While stakeholders may be willing to "trust the experts" if an evaluation offers results that support what they want to hear, there may be a reasonable tendency to distrust results – and particularly methods – that they don't understand.

People tend to trust or distrust evidence based on what they already believe, looking for results that confirm what they believe and looking for ways to discredit contrary information. Perhaps one reason is that it is difficult to distinguish between good and bad evidence. Currently, there is much ongoing work to provide training in measurement and evaluation for donors and policymakers: when individuals have a greater understanding of impact evaluation, they may be better able to recognize differing qualities of evidence, allowing individual evaluations to have greater impact.

Module 9: Evaluation

Effective dissemination

Rapid, broad and well targeted dissemination strategies are important determinants of utilization. One reason that many sound and potentially useful evaluations are never used is that very few people have ever seen them.

Providing rapid feedback to government on issues such as the extent of corruption or other "hot" topics enhances utilization.

Continuous and targeted communication builds interest and confidence and also ensures "no surprises" when the final report and recommendations are submitted. This also allows controversial or sensitive findings to be gradually introduced. Trust and open lines of communication are important confidence builders. An individual evaluation will rarely be entirely conclusive, but conclusions drawn from an accumulation of evidence may be more difficult to refute.

The choice of the institution and the evaluators can contribute to dissemination and credibility of the findings.

Making data available to the academic community is also an important way of broadening interest and support for evaluations and also of legitimizing the methodologies (assuming they stand up to academic critiques as have PROGRESA and Familias en Accion).

Positive and non-threatening findings

Positive evaluations, or those that support the views of key stakeholders, increase the likelihood they will be used. While this is not surprising, one of the reasons is that many agencies are either fearful of the negative consequences of evaluation or considered evaluation as a waste of time (particularly the time of busy managers) or money. Once stakeholders have appreciated that evaluations were not threatening and were actually producing useful findings, agencies have become more willing to request and use

evaluations and gradually to accept negative findings – or even to solicit evaluations to look at areas where programs were not going well.

There is always demand for results that confirm what people want to hear. Concerns over potential negative results, bad publicity, or improper handling of the results may reduce demand; sensitivity, trust-building, and creative arrangements may help overcome these fears. Consequently, there may be some benefit in taking advantage of opportunities to present good results, especially if it helps the process of getting stakeholders to understand and appreciate the role of impact evaluation.

Active engagement with national counterparts

The active involvement of national agencies in identifying the need for an evaluation, commissioning it, and deciding which international consultants to use is central to utilization. It gives ownership of the evaluation to stakeholders and helps ensure the evaluation focuses on important issues. It often increases quality by taking advantage of local knowledge and in several cases reduces costs (an important factor in gaining support) by combining with other ongoing studies.

This cooperation can enable evaluators to modify the initial evaluation design to reflect concerns of clients – for example, changing a politically sensitive randomized design to a strong quasi-experimental design.

Involving a wide range of stakeholders is also an important determinant of utilization. This can be achieved through consultative planning mechanisms, dissemination and ensuring that local as well as national level agencies are consulted.

In some contexts, the involvement of the national statistical agency increases the government's trust, and the results and the process have been better accepted when overseen and presented by the statistics agency.

Module 9: Evaluation

Demonstrating the value of evaluation as a political and policymaking tool and adapting the design to the national and local political contexts

When evaluation is seen as a useful political tool, this greatly enhances utilization. For example, managers or policymakers often welcome specific evidence to respond to critics, support for continued funding or program expansion. Evaluation can also be seen as a way to provide more objective criticism of an unpopular program.

Once the potential uses of planning tools such as cost-effectiveness analysis are understood, this increases the demand for, and use of, evaluations. Evaluations can also demonstrate the practical value of good monitoring data, and increased attention to monitoring in turn generates demand for further evaluations. When evaluations show planners better ways to achieve development objectives, such as ensuring services reach the poor, this increases utilization and influence.

Increasing concerns about corruption or poor service delivery have also been an important factor in government decisions to commission evaluations. In some cases, a new administration wishes to demonstrate its transparency and accountability or to use the evaluation to point out weaknesses in how previous administrations had managed projects. Evaluations that focus on local contextual issues (i.e. that are directly relevant to the work of districts and local agencies) are much more likely to be used.

In cases where a clearly defined selection cut-off point can be defined and implemented (e.g. the score on a poverty or probability of school drop-out scale), the regression discontinuity design (RD) can provide a methodologically strong design while avoiding political and ethical concerns.

Evaluators must adapt evaluation designs to political realities when deciding what evaluation strategies will be both technically sound and politically feasible. Evaluations of large, politically sensitive programs should be designed at an early stage before the programs have developed a large constituency and become resistant to questioning of their goals and

methods. Evaluations should begin early in the program with greater use being made of small pilot projects to assess operational procedures and viability for expansion.

The methodological quality of the evaluation and credibility of the international evaluators

High quality of an evaluation is likely to increase its usefulness and influence. Quality improves the robustness of the findings and their policy implications and may assist in dissemination (especially in terms of publication). However, an impact evaluation of a compromised quality may still be useful if it can provide timely and relevant insight or if it ventures into new territory: new techniques, less-evaluated subject matter, or in a context where relevant stakeholders have less experience with impact evaluations.

The credibility of international evaluators, particularly when they are seen as not tied to funding agencies, can help legitimize high profile evaluations and enhance their utilization. In some cases, the use of what is considered "state of the art" evaluation methods such as randomized control trials can raise the profile of evaluation (and the agencies that use it) and increase utilization.

New and innovative evaluations often attract more interest and support than the repetition of routine evaluations.

On the other hand, while studies on the "frontier" may be more novel or attract more attention, subsequent related studies may be useful in confirming controversial findings and building a body of knowledge that is more accepted than a single study, especially a single study with unpopular findings.

Evaluation methods, in addition to being methodologically sound, must also be understood and accepted by clients. Different stakeholders may have different methodological preferences.

Evaluation capacity development

Evaluation capacity, especially at a local level, is an important factor in the quality of impact evaluations that also affects the ability of stakeholders to demand, understand, trust, and utilize the results.

Capacity building is an iterative process and may improve both demand and quality.

3.3.4. Strategic considerations in promoting the utilization of impact evaluations

Many of the evaluations cited in [3] were selected opportunistically, depending on the availability of donor funding and technical support and the interest of a particular agency, or even a small group of champions within the agency. While individual evaluations may have made a useful contribution, the cases illustrate that the effects and benefits are often cumulative, and utilization and government buy-in tend to increase where there is a sequence of evaluations. In several cases, the first evaluation was methodologically weak (for example, being commissioned late in the project and relying on retrospective data collection methods for reconstructing the baseline), but when the findings were found useful by the national counterparts, this generated demand for subsequent and more rigorous evaluations.

Effective utilization of impact evaluations is an incremental process, with the full benefits only being realized once a number of useful evaluations have been conducted. Policymakers, planners, managers and funding agencies gradually gain confidence in the value of impact evaluation once they have seen some of the practical benefits, and have learned that some of the initial concerns and reservations were not fully justified. **A key element in the successful utilization is developing a system for the selection of evaluations that address key policy issues and for analysis, dissemination, and utilization of the results**. All of these considerations require the *institutionalization of an impact evaluation*

system with strong buy-in from key stakeholders and with a powerful central government champion, usually the ministries of finance or planning.

Institutionalization of Impact Evaluation within the Framework of a Monitoring and Evaluation System (IEG 2008 [16]) identifies a number of different paths towards the institutionalization of impact evaluation and points out that the utility and influence of many methodologically sound evaluations has been limited because they were looked upon as one-off evaluations and did not form part of a systematic strategy for selecting evaluations that addressed priority policy issues or that were linked into national budget and strategic planning. This report argues that methodologically sound and potentially useful impact evaluations do not automatically ensure the development of an evaluation system, and that the creation of such a system requires a strong commitment on the part of government agencies and donors over a long period of time.(Comment: so true but how to assure it)

The present text and [3] corroborate many of the findings of the IEG study. In addition to the recommendations and guidelines presented in the previous sections, the discussion of the evaluation presentations raised the following issues:

It is important identify and support impact evaluations that can provide findings and knowledge that will be useful to a broader audience than the project agency whose programs are being evaluated.

The role of the evaluator should be clarified. Should they become advocates for the adoption of the evaluation findings (for example, the free distribution of anti-malaria or de-worming treatments) or should their role be limited to the collection and analysis of data that the evaluation clients will interpret? While many clients require the evaluator to present recommendations, there is a concern in the evaluation profession that the requirement to present recommendations may lead to a bias in how the findings are presented (and particularly ignoring findings that do not support the recommendations).

There is also a challenge when academics are asked to provide recommendations. The academic researcher is trained to present caveats rather than to come to firm conclusions. Also, the academic has a different set of incentives, and she or he is often judged on the number of publications (in journals that require the use of particular methodologies and give less value to policy recommendations based on the best available, but less rigorous, evidence).

The previous point relates to a concern that the influential role of academic researchers in the program evaluation field means that many evaluations are method-driven rather than policy driven. This criticism has often been levelled at advocates of randomized control trials who are seen as ignoring important policy evaluations where it is not possible to use rigorous methods, in favour of evaluations that are less useful to policymakers and planners but where it is possible to use randomized designs.

Further to this point was the recommendation that there is a need to consider rules and procedures for defining acceptable standards of evidence. Different fields, such as health and drug research, may traditionally use different standards of evidence and proof than those used in other fields such as conditional cash transfers and poverty analysis. Is it possible to define generally accepted standards of evidence that can apply in all sectors?

The question of standards of evidence also applies to increasing use of mixed method evaluation designs that recognize and seek to reconcile the different criteria of evidence and proof conventionally used in quantitative and qualitative research.

A final point concerned the question of whether all evaluation results should be disseminated. For example, if the success of an evaluation depends on close cooperation of national counterpart agencies, should there be situations in which these agencies can decide whether and when certain findings should be disseminated? There are other situations in which potentially important but controversial findings may be based on weak evidence (for example with small sample sizes and low statistical power). While researchers may understand that such findings must be interpreted with caution, the mass media or political supporters or critics of a program may ignore these caveats, perhaps jumping to conclusions

that a program should be terminated or an innovative approach should receive major funding.

Exercises

Select a topic of your working background. If you do not have an example at hand, please select one of the monitoring examples of the documentary (Module 4 & Module 7)

Design an evaluation framework: list the tools, surveys, communication and data gathering strategies necessary to meet your objectives with the different stakeholder groups. Use reference chart in 1.11 (evaluation framework should be given in summary, sketch-like form

Use the Awareness chart 1.12 to identify elements of the intended evaluation in a practical sense, propose tools and strategies to:

Create demand

Strengthen methodology

Strengthening program design

Providing evidence

Providing evidence to challenge by stakeholders

Tools and strategies should be given in summary, sketch-like form to make clear the intended outcomes

Refer to the examples in chapter 3. Which lessons can be learned?

Bibliography

Documents

[1] UNEG, 'Norms for Evaluation in the UN System', 2005, available at: http://www.unevaluation.org/unegnorms ; and UNEG, 'Standards for Evaluation in the UN System', 2005, available at: http://www.unevaluation.org/unegstandards.

[2] World Bank Group: Handbook on Impact Evaluation Quantitative Methods and Practices, Shahidur R. Khandker, Gayatri B. Koolwal, Hussain A. Samad: Handbook Evaluation-WB.pdf

[3] Making Smart Policy: Using Impact Evaluation for Policy Making- World Bank Group June 2009

[4] Has Economic Growth in Mozambique Been Pro-Poor? Discussion Paper 2005 International Food Policy Research Institute Washington D.C: IFPRI_hasDevelopmentbeenProPoor.pdf

[5] Is Poverty Decreasing in Mozambique? Joseph Hanlon, Senior Lecturer,Open University, England 2007

[6] Do bicycles equal development in Mozambique? Von Joseph Hanlon,Teresa Smart Rochester NY and Woodbridge: James Currey, 2008. ISBN 978 184701 319 4.

[7] Project Proposal for a Survey Measuring Development in Sofala on District 2007 LevelKRoeder_ProposingSMD_Sofala070301.pdf

[8] Heterogeneous Impacts of Conditional Cash Transfers: Evidence from Nicaragua Ana C. Dammert - Dammert_Nicaragua_2008.pdf

[9] Information Distortion, Elite Capture, and Task Complexity in Decentralised Development by Jean-Philippe Platteau - Information Distortion, Elite Capture.. in Decentralized Development- Platteau - 2008.pdf

Unit 1: Not yet defined

[10]   Evaluating three stylised interventions Author: Martin Ravallion – in Development Effectiveness - Affiliation: a Development Research Group, World Bank, Washington, DC, USA

[11]   ASTA Wirtschafts- und Sozialstatistisches Archiv Band 4-Heft 4-2011, J.Frohn Glück, Nachhaltigkeit und Ökonometrie Zur Einbeziehung sozialer Aspekte in makroökonomoischer Modelle

[12]   Hatry, Harry P. 1999. Performance Measurement: Getting Results. Washington, D.C.: The Urban Institute Press.

[13]   OECD (Organisation for Economic Co-operation and Development). 2001. "Evaluation Feedback for Effective Learning and Accountability." Paris: OECD/ DAC.

Links

[14]   http://www.google.de/url?sa=t&source=web&cd=1&ved=0CB0QFjAA&url=http%3A%2F%2Fsiteresources.worldbank.org%2FINTPA%2FResources%2FTraining-Materials%2FMozambique_BeatingTheOdds_handout.pdf&rct=j&q=World%20Bank%2C%20Beating%20the%20Odds%3A%20Sustaining%20Inclusion%20in%20a%20Growing%20Economy%20%E2%80%93%20A%20Mozambique%20Poverty%2C%20Gender%20and%20Social%20Assessment%202007.&ei=1z1dTZS9IsfasgahwuHeCg&usg=AFQjCNHw9dL9J3NA7HEqorBFPF1sz1V2VA&sig2=BR2I572_1Qv4NosyAeN7gw

[15]   No link available

[16]   http://books.google.de/books?id=kVs2QiJof_8C&printsec=frontcover&dq=Joseph+Hanlon+and+Teresa+Smart.+Do+Bicycles+Equal+Development+in+Mozambique&source=bl&ots=V6uA2nSVlH&sig=vXSDuG5nJbpxrsY0TmZ50dFyM2w&hl=de&ei=nkxdTaLINYKi8QPhgb2KCw&sa=X&oi=book_result&ct=result&resnum=1&sqi=2&ved=0CBgQ6AEwAA#v=onepage&q&f=false

Unit 1: Not yet defined

[17]   http://www.klaus-

roeder.com/4_Projekte/Trainings_Projects/1987_1990%20Harambee.html

[18]   ftp://ftp.fao.org/docrep/fao/010/i0132e/i0132e03b.pdf

http://lnweb90.worldbank.org/oed/oeddoclib.nsf/DocUNIDViewForJavaSearch/E629534B7C6

77EA78525754700715CB8/$file/inst_ie_framework_me.pdf