



Programme Manager's Planning Monitoring & Evaluation Toolkit

Division for Oversight Services

August 2004¹

Tool Number 5: Planning and Managing an Evaluation Part II: Defining Evaluation Questions and Measurement Standards

I. Introduction

The toolkit is a supplement to the UNFPA programming guidelines. It provides guidance and options for UNFPA Country Office staff to improve planning, monitoring and evaluation (PM&E) activities in the context of results based programme management. It is also useful for programme managers at headquarters and for national programme managers and counterparts.

This part II of tool number 5 discusses the “what” of evaluation: steps to define evaluation questions and measurement standards. The content is based on a review of evaluation literature from bilateral and other development agencies such as such as Danida, ILO, Management Sciences for Health as well as documentation from UNFPA project evaluations.

II. Defining Evaluation Questions

Most evaluations are concerned with issues of programme design, delivery and performance. Design and delivery issues refer to factors affecting results. These factors appear during programme implementation.

Performance issues relate to the actual programme results (see **Box 1**). Each of these issues is explained in greater detail below.

¹This tool was first published in December 2000.

Box 1. What do we mean by Result?

A result is a describable or measurable change in state that is derived from a cause and effect relationship. Results are the effects generated by a programme.

There are *three different types* of results:

Outputs Products and services that result from the completion of activities within a development intervention.

Outcomes The intended or achieved short and medium-term effects of an intervention's outputs, usually requiring the collective effort of partners. Outcomes represent changes in development conditions which occur between the completion of outputs and the achievement of impact.

Impacts Positive and negative long term effects on identifiable population groups produced by a development intervention, directly or indirectly, intended or unintended. These effects can be economic, socio-cultural, institutional, environmental, technological or of other types.

Source: Tool Number 1: Glossary of Planning, Monitoring and Evaluation Terms, March 2004.

Validity of design

A good programme design guides the implementation process, facilitates monitoring of implementation and provides a solid base for performance evaluation. In UNFPA, issues of programme design are assessed by using the programme logical framework.

Some key questions related to design include²:

- **Outputs, outcomes and impact (the results):** are they clearly stated, describing solutions to identified problems and needs?
- **Inputs and strategies:** are they identified and are they realistic, appropriate and adequate to achieve the results?
- **Indicators:** are they direct, objective, practical and adequate (DOPA)³? Is responsibility for tracking them clearly identified?

² These questions are illustrative and should not be used as a "blue print".

³ A **Direct** Indicator closely tracks the result it is intended to measure; an **Objective** Indicator is unambiguous about: 1) what is being measured and data being collected; 2) has a clear operational definition that is independent of the person measuring the indicator; a **Practical** Indicator can be gathered at reasonable cost and frequency, and can be available in time for use in decision-making; an **Adequate** indicator constitutes the minimum necessary to ensure that progress towards results is sufficiently well captured. Further details on indicators are provided in Tool Number 6 Part I: Identifying Output Indicators – The Basic Concepts available at www.unfpa.org

-
- **External factors and risks:** have factors external to the programme that could affect implementation been identified and have the assumptions about such risk factors been validated?
 - **Execution, implementation, monitoring and evaluation responsibilities:** have they been clearly identified?
 - **Gender sensitivity:** does the programme design address the prevailing gender situation? Are the expected gender related changes adequately described in the outputs? Are the identified gender indicators adequate?
 - **Capacity building:** does the programme include strategies to promote national capacity building?
 - **Programme approach:**
 1. In the case of a programme evaluation, does the design clearly establish linkages among programmes components?
 2. In the case of a programme component evaluation, are linkages among its interventions clearly established to ensure synergy in achievement of programme components results?

Delivery process

An assessment of the delivery process focuses on how the programme is being/was implemented to determine if the programme has remained on the right track towards the achievement of planned results and if not, what were the influencing factors.

Some key questions related to the delivery process include:

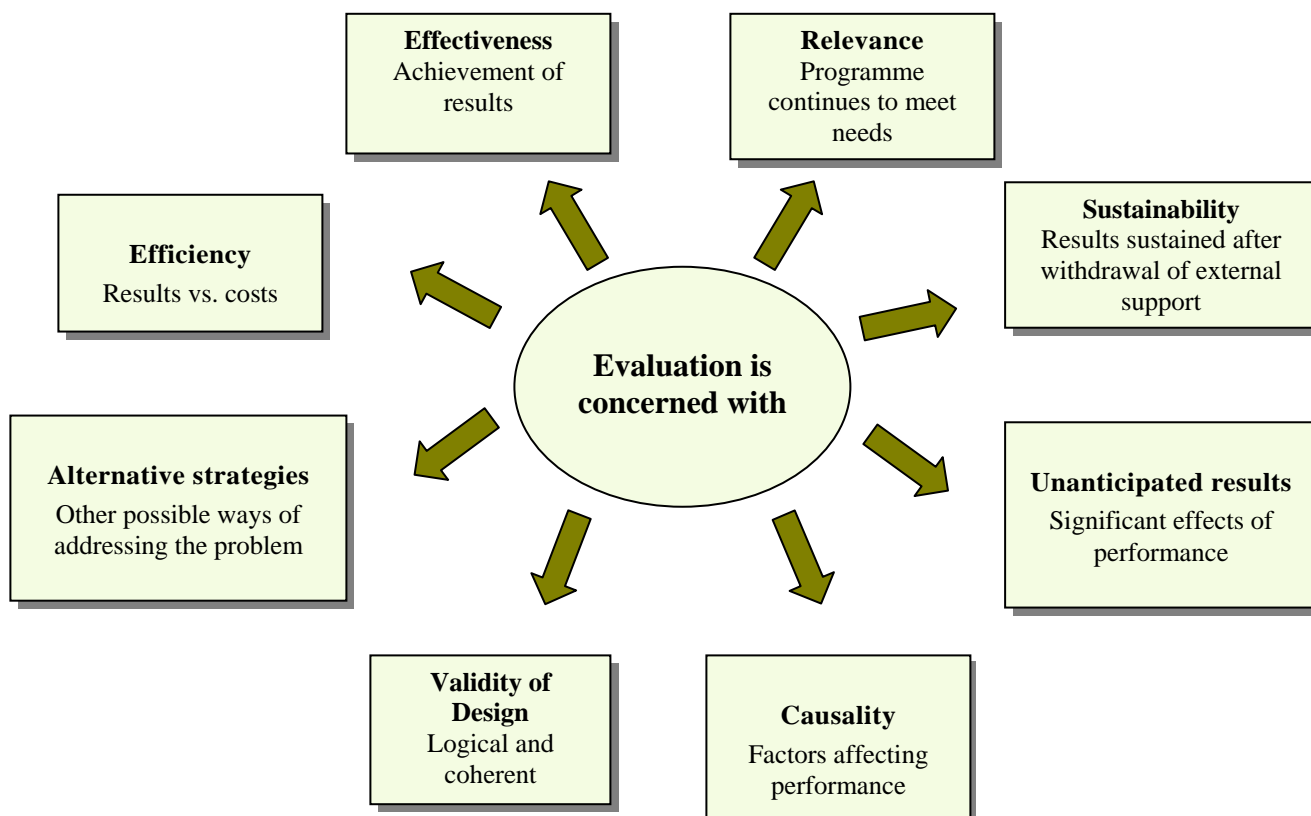
- **Activities:** how were they implemented?
- **Outputs:** were the planned outputs achieved? Were they achieved within the planned time frame? Were they of adequate quality? If not, why?
- **Programme management:**
 1. Did the programme implementers discharge their roles in a cost-effective and cost-efficient manner? If not, why not?
 2. Were sound financial and equipment management procedures practised? Were the financial, human and material resources managed responsibly and efficiently?
 3. Was the technical assistance provided appropriate and of good quality?
 4. Did the monitoring and evaluation systems and processes allow for adequate assessment of changes in risks and opportunities in the internal and external environments? Did they contribute to effective decision-making in the course of programme implementation?

Performance

When assessing programme performance, evaluations look beyond the delivery process and focus on the results of inputs delivered and the work done. The outcome of this assessment determines whether or not the programme has achieved or is likely to achieve its outputs and contribute to achieving programme outcomes and impact.

The core evaluation concerns to assess programme performance are illustrated in **Figure 1** and described below.

Figure 1. Core Evaluation Objectives



Source: ILO, 1997.

Relevance

An assessment of programme relevance examines the appropriateness of results in relation to: the national needs, policies, and priorities; the needs and priorities of programme target groups (the local programme context); UNFPA's policies and priorities and its comparative advantage vis à vis other UN agencies and development partners. The analysis ascertains whether the programme continues to make sense and identifies any changes that may have occurred in its context during implementation. The initial problems and needs may no longer exist and policies and priorities may have changed as a result of political,

economic, social and other factors, or even because of programme activities. Ultimately, the analysis determines whether the results are still valid or should be reformulated.

Some key questions related to relevance include:

- **Needs, mandates, policies and priorities:** Do the programme planned results address the national needs? Are they in line with the government's priorities and policies? Are they in line with UNFPA's mandate? Does the target population consider them useful? Are they complementary to other donor interventions? Should results be adjusted, eliminated or new ones added in light of new needs, priorities and policies?

Effectiveness

An assessment of programme effectiveness focuses on the extent to which the outputs have been or will be achieved and whether the programme is likely to contribute to the stated outcomes and impact. If not, the evaluation will identify whether the results should be modified (in case of a mid-term evaluation) or the programme be extended (in case of a final evaluation) in order to enable achievement of stated results.

Some key questions related to effectiveness include:

- **Outputs:** to what extent have planned outputs been or will be achieved? What is the quality of the outputs?
- **Data on indicators:** have data been collected on the indicators of achievement? Do they provide adequate evidence regarding achievement of programme outputs and contribution to outcomes and impact? Is it necessary to collect additional data?
- **Gender:** what were the achievements in terms of promoting gender equity and equality (planned/unplanned)?
- **Capacity development:** what were the achievements in terms of capacity development (planned/unplanned)?

Efficiency

An assessment of programme efficiency measures the “productivity” of the programme interventions. It assesses the results obtained in relation to the expenditure incurred and resources used by the programme during a given period of time. The analysis focuses on the relationship between the quantity, quality, and timeliness of inputs, including personnel, consultants, travel, training, equipment and miscellaneous costs, and the quantity, quality, and timeliness of the outputs produced and delivered. It ascertains whether there was adequate justification for the expenditure incurred and examines whether the resources were spent as economically as possible.

Some key questions related to efficiency include:

- **Costs:** did the actual or expected outputs justify the costs incurred? Have the resources been spent as economically as possible?
- **Duplication:** did programme activities overlap and duplicate other similar interventions (funded nationally and/or by other donors)?
- **Alternative options:** are there more efficient ways and means of delivering more and better outputs with the available inputs?

Sustainability

An assessment of programme sustainability ascertains the extent to which the programme results have had or are likely to have lasting results after programme termination and the withdrawal of external resources. The factors affecting sustainability are examined on the basis of the priority assigned to the programme by stakeholders. Their readiness to continue supporting or carrying out specific activities, or even replicate the activities in other regions or sectors of the country, is particularly relevant. The analysis also assesses the availability of local management, financial and human resources that would be needed to maintain the programme results in the long run.

Some key questions related to sustainability include:

- **Likely sustainability:** is it likely that programme achievements will be sustained after the withdrawal of external support? Are involved counterparts willing and able to continue programme activities on their own? Have programme activities been integrated into current practices of counterpart institutions and/or the target population?
- **Resources:** have they been allocated by the counterparts to continue programme activities?

Causality

An assessment of causality examines the factors or events that have affected the programme results. If the inputs needed to carry out the planned activities and deliver the expected outputs were available on time, the implementation and performance would be successful. If, on the other hand, there were significant deviations from the planned schedules, the analysis would determine the reasons for such changes. The assessment should also analyse the effect of other factors such as technical, administrative or managerial constraints, inadequate inputs, failed commitment by programme counterparts, insufficient funds, faulty assumptions or the effect of unexpected external factors.

Some key questions related to causality include:

- **What factors:** what particular factors or events have affected the programme results?
- **Internal/external factors:** were these factors internal or external to the programme?

Unanticipated results

A programme evaluation may find significant unforeseen positive or negative results of programme activities. Once identified, appropriate action can be taken to enhance or mitigate them for a greater overall impact.

Some key questions related to unanticipated results include:

- **Were there any** unexpected positive and/or negative results of the programme?
- **How to address them:** can they be either enhanced or mitigated to achieve the desired impact?

Alternative strategies

Evaluations examine whether alternative approaches might have had greater impact or might have been more cost-effective, particularly if the original strategies turn out to be inappropriate. This analysis is especially valuable when follow-up programmes are planned.

Some key questions related to alternative strategies include:

- **More effective approaches:** is there, or would there have been, a more effective way of addressing the problem(s) and satisfying the needs in order to achieve the outputs and contribute to higher level aims?
- **Relevance:** are programme strategies still valid or should they be reformulated?

Not all of the above evaluation objectives have to be examined in every evaluation. The final choice will depend on the purpose of each evaluation⁴. For instance, a formative evaluation undertaken in the course of programme implementation with the aim of taking decisions to improve its design and/or implementation would typically emphasise concerns of design, delivery process, efficiency, causality, unanticipated results, and alternative strategies.

A summative evaluation, undertaken at the end of programme implementation to judge its effectiveness, would typically concentrate on concerns of relevance, effectiveness, efficiency, alternative strategies and sustainability.

An evaluation, which aims at extracting lessons learned and best practices or defining policy options would assess design, delivery processes, causality and efficiency in order to extract those characteristics which can effectively and efficiently deliver the desired results.

⁴ For a discussion on the issue of evaluation purpose, consult Tool Number 3: Purposes of Evaluation.

III. Methodological Challenges

Evaluator(s) face a number of methodological challenges with respect to the standards⁵ they use to measure relevance, effectiveness, efficiency and sustainability. These standards and methodological challenges are summarized in **Table 1** and further discussed below.

Table 1. Performance related Evaluation Objectives: measurement standards and methodological challenges.

Evaluation Objective	Measurement Standards	Methodological Challenge
Relevance	Needs, priorities and policies of programme target population, counterparts; UNFPA's policies, priorities, comparative advantage.	Lack of consensus on or incorrect assessment of needs and country priorities and lack of clear policies. Incorrect assessment of and/or lack of consensus on UNFPA's comparative advantage.
Effectiveness	Agreed outputs, outcomes and impact. Status of affected institutions, target population, and infrastructure prior to the programme interventions.	Unclear, multiple, confusing or changing results statements. Poorly defined results indicators. Lack of baseline information on the affected institutions, people, infrastructure. Poor knowledge of cause /effect linkages. Difficulty in attributing results to the particular programme due to intervening variables.
Efficiency	Similar interventions/best practices; criteria for what is considered reasonable.	What standards to use as a reference.
Sustainability	Sustainability factors (see Box 4)	Long term sustainability is a hypothetical, projected situation. Not all intervening factors which can compromise sustainability can be foreseen.

Source: Adapted from Danida, 1999.

⁵ A standard is a level of performance according to specified criteria or achievement dimensions.

With respect to the *relevance* of programme strategies, it requires in-depth field analysis by evaluator(s) to adequately determine the continued relevance of the planned programme results if the context (needs, priorities and policies) were not clearly defined at the time of programme design or in the course of implementation. **Box 2** provides an example of such a situation from a UNFPA commissioned evaluation.

Box 2. The importance of using contextual information for programme design and adjustment.

The report of an evaluation of the Jenin Community-Based RH project in Palestine concluded the following:

“The second objective stated in the project proposal was that “18,000 new users of family planning will have been recruited and continuation rates will be improved.”

The problem with this objective is twofold. First of all, the target population was overestimated and the target recruitment numbers were too ambitious. The number of target population does not appear to be based on the available scientific evidence (estimates of the district population published by the Palestinian Central Bureau of Statistics). Secondly, the manner in which the numbers were determined is not clear from the proposal document. Furthermore, the project staff did not seem to be aware of this target and were unable to explain it. Still, the unrealistically high expectations seemed to put pressure on the staff to generate high numbers of new users and to reflect the numbers in their reported statistics, to the exclusion of other important activities that were being carried out.

As for the objective of improving continuation rates, it was not clear how that would be verified, since no baseline figure for continuation rates in those communities existed.

As a consequence of lack of proper initial situation analysis and adjustment in the course of project implementation, a large proportion of the budget (80,000 USD) was allotted to the purchase of contraceptives many of which eventually expired on the shelf due to lack of demand.”

Source: Halabi, January 2000.

Standards to determine a programme’s *effectiveness* often have to be reconstructed by evaluator(s) when planned programme results are too ambitious in relation to the resources and time frame available. Additionally, the programme indicators are frequently poorly or incorrectly defined thus hampering a sound assessment of achievement of programme outputs. The failure of programme implementers to gather baseline data at the beginning of the implementation process against which progress can be measured, constrains the evaluator(s) ability to assess results. To facilitate the objective evaluation of results achievement, the programme indicators, particularly those related to outputs, should be adjusted and refined in the early phase of programme implementation based on collection of baseline data and the accumulated knowledge of the programme context.

An additional difficulty is that long-term results can usually only be determined with certainty a significant period of time after programme completion. During that time, developments external to the programme such as economic and social development factors, (for instance increase in age at marriage) could have influenced the programme targets thereby making it difficult to ascribe improvements to the programme interventions.

Defining objective *efficiency* standards is a major challenge for evaluators of UNFPA's programmes and projects. In practice the evaluator(s) frequently rely on their expert judgment, which can be subjective. However, approaches are available to define standards, among others in the health field. The Continuous Quality Improvement tool to strengthen FP programmes is one such approach, which, if used in the course of programme implementation, greatly facilitates monitoring and evaluation of programme efficiency (see **Box 3**). Another good approach for identifying efficiency standards is "benchmarking", analysing the performance of organizations, which excel in areas of work relevant to UNFPA.

Box 3. Using the Continuous Quality Improvement (CQI) approach to define efficiency standards

CQI is performed by teams of staff at RH service delivery points. The CQI team implements the 7 step CQI cycle:

- Step 1:** Identify an area where opportunities for improvement exist
- Step 2:** Define a problem within that area, and outline the sequence of activities (the process) that occurs in that problem area
- Step 3:** Establish the desired outcomes of the process and the requirements needed to achieve them
- Step 4:** Select specific steps in the process to study and for each step, list the factors that prevent the achievement of the desired outcome
- Step 5:** Collect and analyze data about the factors that are preventing the achievement of the desired outcomes of the specific step being studied, and quantify the outcomes of that step
- Step 6:** Take corrective action to improve the process
- Step 7:** Monitor the results of the actions taken.

In **step 3** the CQI team defines the standards of efficiency against which services will be monitored and evaluated. The following is an example of such a standard:

"The client registration process is completed within 30 minutes of client's arrival at the clinic"

Source: The Family Planning Manager, Volume II, Number 1, January/February 1993.

Long-term programme **sustainability** is hard to foresee as many factors intervene over time. For instance, governments change and so may policies that are critical to support certain programmes originally funded by UNFPA. With a government change, key administrators also change and with them the institutional memory so necessary to keep particular approaches and programmes running. A severe economic crisis may appear, jeopardizing funding for the programme. However, programme designers must ensure that the sustainability factors listed in **Box 4** are fully considered at the time of situation analysis and programme design. Evaluator(s) assess the likelihood of sustaining programme activities using the same standards.

Box 4. Factors which influence sustainability of programme activities

1. **Policy Support Measures**, priorities and commitments of programme implementers and target groups.
2. **Choice of Technology** (for instance contraceptives) is appropriate to existing socio-cultural and economic conditions.
3. **Environmental aspects** such as management of population growth and distribution in relation to available land, water, fuel. Management of their living conditions such as housing, waste disposal, drinking water supply in order to avoid epidemics.
4. **Socio-cultural** integration. Acceptance of interventions because they are consistent with local traditions of groups (gender, ethnic, religious).
5. **Organizational** capacity to manage programme activities.
6. **Economic** viability and financial sustainability.

Source: Danida, 1999.

Sources

Halabi, Hanan; Salem, Ruwaida; Wick, Laura. **“Jenin Community Based RH Education Project”**, Project-end Evaluation for UNFPA, Birzeit University, Institute of Community and Public Health, January 2000.

ILO. **“Guidelines for the Preparation of Independent Evaluations of ILO Programmes and Projects”**, Evaluation Unit, Bureau of Programming and Management, November 1997. Available in English, French and Spanish on the web at:

<http://www.ilo.org/public/english/bureau/program/guides/evalmenu.htm>

Management Sciences for Health. **“The Family Planning Manager”**, Volume II, Number 1, January/February 1993. The full Health Manager’s Toolkit is available in English, French and Spanish on the web at <http://erc.msh.org/index.cfm>

Ministry of Foreign Affairs, Danida. **“Evaluation Guidelines”**, February 1999.

This tool is subject to constant improvement. We welcome any comments and suggestions you may have on its content. We also encourage you to send us information on experiences from UNFPA funded and other population programmes which can illustrate the issues addressed by this tool. Please send your inputs to:

**United Nations Population Fund
Division for Oversight Services**

**Daily News Building
220 East 42nd Street
New York, NY 10017**

Telephone: (212) 297-5213

Fax: (212) 297-4938

E-mail: dos@unfpa.org

This tool is posted on the UNFPA website at <http://www.unfpa.org/>