Chapter 5. MONITORING AND EVALUATION

5.1. Monitoring and evaluation (M&E) are integral and individually distinct parts of programme preparation and implementation. They are critical tools for forward-looking strategic positioning, organisational learning and for sound management.

5.2. This chapter provides an overview of key concepts, and details the monitoring and evaluation responsibilities of Country Offices, Regional Offices and others. While this and preceding chapters focus on basic description of monitoring and evaluation activities that CO are expected to undertake, more detailed explanation on practical aspects of managing monitoring and evaluation activities can be found in the UNICEF Monitoring and Evaluation Training Resource as well as in the series Evaluation Technical Notes.

5.3. M&E concepts and accountabilities apply equally in developmental and humanitarian settings. While important differences are noted where appropriate, not all eventualities can be foreseen. Creative adaptation to the particular context will always be needed, including to the special stresses and requirements of crisis settings.

Section 1. Key Conceptual Issues

5.4. As a basis for understanding monitoring and evaluation responsibilities in programming, this section provides an overview of general concepts, clarifies definitions and explains UNICEF’s position on the current evolution of concepts, as necessary.
5.5. Inspection, audit, monitoring, evaluation and research functions are understood as complementary but very distinct management tools. They can be situated along a scale, as in Figure 5.1. At one extreme, inspection and audit can best be understood as control functions. At the other research and to an extent evaluation are meant to generate knowledge. Performance monitoring, some elements of audit, and evaluation are tools to facilitate results-based management. The distinctions are elaborated in the sub-sections below.

5.6. The linkages of evaluation and monitoring with successful results-based management cannot be over-emphasized. Both monitoring and evaluation are meant to permit more effective decision-making, including decisions to improve, reorient or discontinue the evaluated intervention or policy; decisions about wider organisational strategies or management structures; and decisions by national and international policy makers and funding agencies. Monitoring and evaluation create a valid evidence base for making informed programming decisions. Section 5.1.2 below expands on these concepts further.

**Audits**

5.7. Audits generally assess the soundness, adequacy and application of systems, procedures and related internal controls. Audits encompass compliance of resource transactions, analysis of the operational efficiency and economy with which resources are used and the analysis of the management of programmes and programme activities. See [CF/EXD/2005-004](#) for the *Charter of Authorities and Responsibilities of the Office of Internal Audit*, and [E/ICEF/2003/AB/L.11](#) on Internal Audit activities).

5.8. At country level, Programme Audits may identify the major internal and external risks to the achievement of the programme results, and weigh the effectiveness of the actions taken by the UNICEF Representative and CMT to manage those risks and maximise programme achievements. These are important components of results-based management. While audits may overlap somewhat with evaluation, they do not generally examine the relevance or impact of a programme, which is a key role of evaluation as described below. A Programme Management *Audit Self-Assessment Tool* is contained in Chapter 6.
Monitoring

5.9. There are two kinds of Monitoring:

- **Situation monitoring** measures change in a condition or a set of conditions or lack of change. Monitoring the situation of children and women and goals such as the MDGs is necessary when trying to draw conclusions about the impact of programmes or policies. It also includes monitoring of the wider context, such as early warning monitoring, or monitoring of socio-economic trends and the country’s wider policy, economic or institutional context. UNICEF is broadly engaged in situation monitoring using the CCA and Situation Analysis, DevInfo, and MICS among other tools.

- **Performance monitoring** measures progress in achieving specific results in relation to an implementation plan, whether for programmes, strategies, or activities. It is core accountability for effective work planning and review.

5.10. These two types of monitoring feed into two types of evaluations. Impact evaluations measure the results of programmes or policies on children, their families, and their communities. Situation monitoring tools provide the measures of change in children’s lives; thus, situation monitoring is a critical input to effective impact evaluation, including for the Millennium agenda and its goals. Process evaluations measure the quality and efficiency of the programme –whether it is well organized, meeting its activity and output targets, adequately participatory, etc.
5.11. Evaluation is an exercise that attempts to determine as systematically and objectively as possible the worth or significance of an intervention, strategy or policy. The appraisal of worth or significance is guided by key criteria discussed below. Evaluation findings should be credible, and be able to influence decision-making by programme partners on the basis of lessons learned. For the evaluation process to be ‘objective’, it needs to achieve a balanced analysis, recognise bias and reconcile perspectives of different stakeholders (including primary stakeholders) through the use of different sources and methods.

5.12. An evaluation report should include the following:

- Findings – factual statements that include description and measurement;
- Conclusions – corresponding to the synthesis and analysis of findings;
- Recommendations – what should be done, in the future and in a specific situation; and, where possible,
- Lessons learned – corresponding to conclusions that can be generalised beyond the specific case, including lessons that are of broad relevance within the country, regionally, or globally to UNICEF or the international community. Lessons
can include generalised conclusions about causal relations (what happens) and generalised normative conclusions (how an intervention should be carried out). Lessons can also be generated through other, less formal evaluative activities.

5.13. It is important to note that many reviews are in effect evaluations, providing an assessment of worth or significance, using evaluation criteria and yielding recommendations and lessons. An example of this would be a Mid Term Review of the UNICEF-supported Country Programme (discussed in detail in Chapter 4). Country offices should employ evaluation concepts and approaches during reviews as well as when evaluations are discrete activities.

Research and studies

5.14. Research, studies and evaluations draw on the same sets of methodologies and thus have the same concerns for independence, rigour, and quality standards. All must provide valid evidence for the decisions they will influence. Choices of scope, model, methods, process and degree of precision must be consistent with the questions that the evaluation, study or research is intending to answer.

5.15. In the simplest terms, an evaluation focuses on a particular intervention (project, programme, strategy, or broad policy) or set of interventions, and results in an analysis of performance or impact compared to some measure of effectiveness. An evaluation explains as much as possible why the changes have occurred. It is also relevant to pitch a level of higher-quality analysis than what is normally possible to do, e.g. support to the recent series of articles in the Lancet on child development and child nutrition, and others.

5.16. Research and studies are targeted efforts to develop new knowledge. They tend to address a broader range of questions, often dealing with socio-economic conditions and causal analyses that serve as a reference for programme design. A Situation Analysis or CCA thus fall within the broader category of “research and study”.

5.17. Research, like monitoring, is closely linked at times to evaluation. For example, baseline surveys provide the measurement base for an evaluation to determine impact. Also very important is “operational” or “action-oriented” research to test parts of the programme design. It often takes the form of intervention trials (e.g. Approaches to Caring for Children Orphaned by AIDS and other Vulnerable Children – Comparing six Models of Orphans Care, South Africa 2001). More generally, pilot
projects are a combination of targeted research and evaluation (See also Chapter 6, Section 18 on Piloting). While not a substitute for evaluation, such research can be useful for improving programme design and implementing modalities.

**Purposes of monitoring and evaluation**

5.18. There are 5 essential purposes that M&E fulfils at regional, national and sub-national levels. These vary according to programme content, counterpart interests, among many other factors.

**Informing Decision Making**

5.19. By gathering and assessing information on organizational and programme performance, M&E permits managers to base their decisions on relevant, valid, comprehensive, and up-to-date data. This is a need at all levels of programming and is particularly important during review exercises and programme design processes.

**National and global learning**

5.20. M&E efforts based good programme design ultimately answer the questions “What works?” and “What does not work?” Understanding what works and ensuring that the lessons learned are disseminated to national and global knowledge networks (internal and external) helps accelerate learning, avoid error and improve efficiency. It is important to harvest the rich evidence base, particularly resulting from innovative programming areas where the quality of the evidence about what works is still too little and/or uncertain. For example, the growth in programme cooperation for Child Protection presents opportunities for well executed evaluations to influence programme strategy choices at the global level.

5.21. The value of lessons learned from good M&E practices goes beyond programme design. These lessons provide the basis for broader advocacy to strengthen policies and programmes through impartial and credible evidence. Evaluations of successful pilot projects provide the necessary rigor to advocate for scaling-up with partners and/or to leverage other investments or key policy changes.

**Accountability**

5.22. M&E also serve accountability purposes. Performance monitoring helps to establish whether accountabilities are met for implementing a programme plan. Evaluation helps to assess whether accountabilities
are met for expected programme results. Global monitoring of the situation of children and women assists in assessing whether national and international actors are fulfilling their obligations and commitments in ensuring the realisation of human rights and the Millennium Declaration/MDGs.

5.23. As a result of M&E-provided analysis, decision makers may make choices that shift resources, close down or expand programmes. It may also influence adjustments in staffing levels and duties, and determine where UNICEF and others have comparative advantages for new programme areas.

5.24. Increasingly, evaluation processes are used that foster wider participation, ensure dialogue, build consensus, and create “buy-in” for recommendations. Involving stakeholders and rights-holders in M&E activities can be tremendously empowering, imparting skills, information, and self-confidence. It can also help participants become effective advocates and learners. Employing M&E as a programming strategy to achieve empowerment can be very effective, and can also support the other purposes of M&E.

**Capacity development**

5.25. National partners expect that their national M&E capacities be strengthened as a result of the UNICEF presence; independent of the value that M&E brings to programme management and learning needs. As with empowerment, M&E can be a programming strategy with specific goals for developing sustainable national capacity. UNICEF has a long history in this area, including building sectoral information systems, DevInfo, MICS, national evaluation associations, Social Observatories and other innovations.

**Evaluation criteria and the issue of attribution**

5.26. In recent years, the evaluation profession has defined more clearly its role and values, especially in international development programming. These now guide all of UNICEF’s evaluation efforts. They are specific enough that staff can employ them in designing and overseeing evaluations. Familiarity with them is a basic accountability of all staff with responsibilities in this area.

**Evaluation criteria**

5.27. The most specific set of guiding criteria are those of the OECD-DAC countries. Adopted by UNICEF and almost all other development actors in the 1990s, the standard OECD-DAC evaluation criteria should guide the appraisal of any intervention or policy. They are:
Relevance

- What is the value of the intervention in relation to other primary stakeholders' needs, national priorities, and national and international partners' policies (including the Millennium Development Goals, National Development Plans, UNDAF, PRS and SWAps)?
- What is the value of the intervention in relation to global references such as human rights, humanitarian law and humanitarian principles, the CRC and CEDAW? For UNICEF, in particular:
- What is the relevance in relation to the MTSP, the CCCs, and foundation strategies – the Human Rights-based Approach to Programming, Gender Mainstreaming and Results-based Management? These global standards serve as a reference in evaluating both the processes through which results are achieved and the results themselves, be they intended or unintended.

Efficiency: Does the programme use the resources in the most economical manner to achieve its objectives?

Effectiveness: Is the activity achieving satisfactory results in relation to stated objectives?

Impact:

- What are the results of the intervention - intended and unintended, positive and negative - including the social, economic, environmental effects?
- How do the results affect the rights and responsibilities of individuals, communities and institutions?

Sustainability:

- Are the activities and their impact likely to continue when external support is withdrawn?
- Will the strategy be more widely replicated or adapted? Is it likely to go to scale?

5.28. For any individual evaluation, the criteria must be converted into specific evaluation questions. The list above is the generic formulation. During the evaluation design phase, it is critical to develop the precise questions to be answered, and to verify that those answers will measure impact etc.

5.29. Evaluation operates effectively within a results
based management approach, such as that used by UNICEF. Figure 5.2 offers a visual illustration of this link. An evaluation planning process that carefully identifies objectives and indicators which meet the evaluation criteria and are founded on results based programming logic will lead to a very strong evaluation.

5.30. The criteria listed above apply equally to developmental and humanitarian action settings. In addition, evaluations of humanitarian action should be guided by four additional criteria as outlined in OECD-DAC guidance:

Coverage:
- Which groups have been reached by a programme?
- What is the differential impact on those groups?

Coordination: What are the effects of co-ordination / lack of co-ordination on humanitarian action?
**Coherence**

- Is there coherence across policies guiding the different actors in, for example the security, developmental, trade, military and humanitarian spheres?
- Are humanitarian considerations taken explicitly into account by these policies?

**Protection**: Is the response adequate in terms of the protection of different groups?

5.31. More detail on these evaluation criteria is provided in the [Evaluation Technical Notes](#). This includes additional evaluation criteria that apply specifically to humanitarian crisis situations. Broader standards that guide the entire evaluation function are discussed in Section 5.3.

**Attribution and partnership**

5.32. As defined by OECD-DAC, attribution represents "the extent to which observed development effects can be attributed to a specific intervention or to the performance of one or more partners taking account of other interventions, (anticipated or unanticipated) confounding factors, or external shocks." For UNICEF, the challenge is to draw conclusions on the cause-and-effect relationship between programmes/projects and the evolving situation of children and women. It may be difficult to attribute intermediate and long-term results to any single intervention or actor. Evaluations and reporting on results should therefore focus on plausible attribution or credible association.

5.33. Difficulties in attribution to any one actor increase as programmes succeed in building national capacity development and sector-wide partnerships. In such cases, it may be sensible to undertake joint evaluations, which may plausibly attribute wider development results to the joint efforts of all participating actors. Multi-agency evaluations of effectiveness of SWAps and CAPs, or the UNDAF Evaluation, are possible examples.