CHAPTER 5. MONITORING AND EVALUATION

SECTION 1. KEY CONCEPTUAL ISSUES

Situating monitoring and evaluation as oversight mechanisms
Evaluation criteria
Purpose of monitoring and evaluation

SECTION 2. SITUATING EVALUATIVE ACTIVITIES IN THE PROGRAMME PROCESS

Monitoring and Evaluation in Emergencies

SECTION 3. MONITORING AND EVALUATION RESPONSIBILITIES IN UNICEF

Integrated Monitoring, Evaluation and Research Plan (IMEP)
Quality standards
Management of monitoring and evaluation resources
Disclosure
Management of effective learning

PQAA CHECKLIST – MONITORING AND EVALUATION

REFERENCES AND RECOMMENDED READING – MONITORING AND EVALUATION

RELATED PPPM REFERENCES

ADDITIONAL REFERENCES

Report on the Evaluation Function in the Context of the Medium-Term Strategic Plan

UNEG Norms for Evaluation in the UN System

UNEG Standards for Evaluation in the UN System

Programme Evaluation Standards

UNICEF Evaluation Report Standards

Children Participating in Research, Monitoring and Evaluation (M&E) — Ethics and Your Responsibilities as Manager - Evaluation Technical Notes 1

UNICEF Standards for preparing the Terms of Reference – Evaluation Technicate Notes 2

Writing a Good Executive Summary – Evaluation Technical Notes 3

Revised July 23, 2005
Rema Venu, Evaluation Office, NYHQ
Chapter 5. MONITORING AND EVALUATION

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.

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.

Section 1. Key Conceptual Issues

3. 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.

Situating monitoring and evaluation as oversight mechanisms

4. Both monitoring and evaluation are meant to influence 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.

5. Inspection, audit, monitoring, evaluation and research functions are understood as different oversight activities situated along a scale (see Figure 5.1). At one extreme, inspection can best be understood as a control function. At the other extreme, research is meant to generate knowledge. Country Programme performance monitoring and evaluation are situated in the middle. While all activities represented in Diagram 5.1 are inter-related, it is also important to see the distinctions.

Monitoring

6. 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 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.
   - **Performance monitoring** measures progress in achieving specific results in relation to an implementation plan, whether for programmes, strategies, or activities.
Evaluation

7. 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.

8. 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.

9. 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 is the Mid-Term Review of the UNICEF-supported Country Programme.
**Audits**

10. 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.

11. At country level, Programme Audits may identify the major internal and external risks to the achievement of the programme objectives, and weigh the effectiveness of the actions taken by the UNICEF Representative and CMT to manage those risks and maximise programme achievements. Thus they may overlap somewhat with evaluation. However they do not generally examine the relevance or impact of a programme. A Programme Management Audit Self-Assessment Tool is contained in Chapter 6.

**Research and studies**

12. There is no clear separating line between research, studies and evaluations. All must meet quality standards. 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.

13. In the simplest terms, an evaluation focuses on a particular intervention or set of interventions, and culminates in an analysis and recommendations specific to the evaluated intervention(s). Research and studies tend to address a broader range of questions – sometimes dealing with conditions or causal factors outside of the programme – but should still serve as a reference for programme design. A Situation Analysis or CCA thus fall within the broader category of "research and study".

14. "Operational" or "action-oriented" research helps to provide background information, or 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). See also Chapter 6, Section 16 on Piloting. While not a substitute for evaluation, such research can be useful for improving programme design and implementing modalities.

**Evaluation criteria**

15. A set of widely shared evaluation criteria should guide the appraisal of any intervention or policy (see Figure 5.2). These are standard OECD-DAC evaluation criteria and have been adopted by UNICEF since 1990:

- Relevance – What is the value of the intervention in relation to other primary stakeholders' needs, national priorities, national and international partners' policies (including the Millennium Development Goals, National Development Plans, UNDAF, PRS and SWAPs), and global references such as human rights, humanitarian law and humanitarian principles, the CRC and CEDAW? For UNICEF, what is the relevance in relation to the
MTSP, the CCCs, and key strategies -- Human Rights-based Approach to Programming 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 on individuals, communities and institutions?
- Sustainability – Are the activities and their impact likely to continue when external support is withdrawn, and will it be more widely replicated or adapted?

16. The evaluation of humanitarian action must be guided by additional criteria as outlined in **OECD-DAC guidance**:

- Coverage - Which groups have been reached by a programme and what is the different 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 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 protection of different groups?

<table>
<thead>
<tr>
<th>GOAL/ INTENDED IMPACT</th>
<th>IMPACT</th>
<th>RELEVANCE</th>
<th>SUSTAINABILITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved health</td>
<td>Intended Reduction in water related diseases Increased working capacity</td>
<td>Whether people still regard water/ hygiene top priority compared with e.g. irrigation for food production</td>
<td>People's resources, motivation, and ability to maintain facilities and improved hygiene in the future</td>
</tr>
<tr>
<td>Objective/ INTENDED OUTCOME</td>
<td>Unintended Conflicts regarding ownership of wells</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improved hygiene</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 5.2 Evaluation Criteria in relation to programme logic**

<table>
<thead>
<tr>
<th>OUTPUTS</th>
<th>EFFICIENCY</th>
<th>EFFECTIVE -NESS</th>
<th>IMPACT</th>
<th>RELEVANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water supplies</td>
<td># of latrines, # of campaigns in relation to plans</td>
<td>Water consumption</td>
<td>Intended Reduction in water related diseases Increased working capacity</td>
<td>Whether people still regard water/ hygiene top priority compared with e.g. irrigation for food production</td>
</tr>
<tr>
<td>Demo latrines</td>
<td>Quality of outputs</td>
<td>Latrines in use</td>
<td>Unintended Conflicts regarding ownership of wells</td>
<td></td>
</tr>
<tr>
<td>Health campaigns</td>
<td>Costs per unit compared with standard</td>
<td>Understanding of hygiene</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

16. The evaluation of humanitarian action must be guided by additional criteria as outlined in **OECD-DAC guidance**:

- Coverage - Which groups have been reached by a programme and what is the different 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 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 protection of different groups?
17. More detail on these evaluation criteria is provided in the Evaluation Technical Notes.

**Purpose of monitoring and evaluation**

*Learning and accountability*

18. Learning and accountability are two primary purposes of monitoring and evaluation. The two purposes are often posed in opposition. Participation and dialogue are required for wider learning, while independent external evaluation is often considered a prerequisite for accountability. On the two extremes, their design – models, process, methods, and types of information – may indeed differ. However, as seen above in Figure 5.1, evaluation sits between these extremes. The current focus on wider participation by internal and external stakeholders and on impartiality allows learning and accountability purposes to be balanced.

19. *Performance monitoring* contributes to learning more locally, ideally at the level at which data are collected and at levels of programme management. It feeds into short-term adjustments to programmes, primarily in relation to implementation modalities. *Evaluation and monitoring of the situation of children and women* contribute to wider knowledge acquisition within the country or the organisational context. *Programme evaluation* not only contributes to improvements in implementation methods, but also to significant changes in programme design. *Evaluation* contributes to learning through both the process and the final product or evaluation report. Increasingly, evaluation processes are used that foster wider participation, allow dialogue, build consensus, and create “buy-in” on recommendations.

20. Monitoring and evaluation also both serve *accountability purposes*. Performance monitoring helps to establish whether accountabilities are met for implementing a 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 commitments in ensuring the realisation of human rights and the Millennium Declaration/MDGs.

*Advocacy*

21. Monitoring and evaluation in UNICEF assisted programmes provide the basis for broader *advocacy* to strengthen global and national policies and programmes for children’s and women’s rights, through providing impartial and credible evidence. Evaluations of successful pilot projects provide the necessary rigour to advocate for scaling-up. Monitoring, particularly situation monitoring, draws attention to emerging children’s and women’s rights issues.

*Early Warning Monitoring Systems*

22. Country Offices should, within the UNCT, assist national governments to establish and operate a basic *Early Warning System* (EWS) and to strengthen the focus of existing systems on children and women. Early warning indicators help to monitor the likelihood of the occurrence of hazards, which have been identified during the preparation of the emergency profile of the Emergency Preparedness and Response Plan (EPRP) (see Chapter 6, Section 8). The most
advanced EWS are presently related to household food security, environmental patterns affecting food production and imminent food crises. These include, for example, the USAID-supported Famine Early Warning System (FEWS), the World Food Programme's Vulnerability Assessment and Mapping System (VAM) and its corresponding Risk Mapping Project (RMP), and the FAO-supported Global Information and Early Warning Systems on Food and Agriculture (GIEWS). One of the key criteria for Early Warning indicators is sensitivity, i.e. that indicators reflect change in the situation promptly. Many such indicators draw on qualitative assessments and non-standardised information systems. Given the different expertise of development partners with such systems, national and sub-national Early Warning Systems should be supported jointly by the UN Country Team, where required.

**Attribution and partnership**

23. 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*.

24. Difficulties in attribution to any one actor increase as programmes succeed in building national capacity building 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 SWAs and CAPs, or the UNDAF Evaluation, are possible examples.
Section 2. Situating Evaluative Activities in the Programme Process

25. There are three groups of evaluation activities, related to different levels of programme management. Each group of activities should guide managers at the corresponding level.

<table>
<thead>
<tr>
<th>Focus</th>
<th>Monitoring activities/systems</th>
<th>Evaluation activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Policy,</td>
<td>MTSP Monitoring</td>
<td>Global, Regional Thematic Evaluations</td>
</tr>
<tr>
<td>Global Strategy,</td>
<td>MDG/Child Goals Monitoring</td>
<td>Global, Regional Syntheses of Evaluations</td>
</tr>
<tr>
<td>Regional Priorities</td>
<td>DevInfo</td>
<td>Meta-Evaluations</td>
</tr>
<tr>
<td></td>
<td>Regional Quality Assurance Systems</td>
<td>Regional Analysis Reports</td>
</tr>
<tr>
<td></td>
<td>Global, Regional Thematic Evaluations</td>
<td>Multi-Country Evaluations</td>
</tr>
<tr>
<td>Country Programme</td>
<td>Situation Assessment and Analysis</td>
<td>Country Programme Evaluation</td>
</tr>
<tr>
<td></td>
<td>Common Country Assessment</td>
<td>Mid-Term Review</td>
</tr>
<tr>
<td></td>
<td>Early Warning Monitoring</td>
<td>Self-Assessment</td>
</tr>
<tr>
<td></td>
<td>Annual Reviews</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Annual Management Reviews</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mid-Term Management Review</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CO Quality Assurance Indicators</td>
<td></td>
</tr>
<tr>
<td>Programme, Programme</td>
<td>Mid-year progress reviews</td>
<td>Programme or project evaluation</td>
</tr>
<tr>
<td>Component</td>
<td>Field visits</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Expenditure tracking (ProMS)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Supply tracking systems</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Annual Management Review</td>
<td></td>
</tr>
</tbody>
</table>

26. When evaluative activities focus on Country Programme strategies and the corresponding choice of interventions, it is important to distinguish between “catalytic” and “operational” programme interventions as highlighted in the MTSP.

27. Different evaluative activities should be situated in relation to CO accountabilities as outlined in Chapter 2 (see Figure 2.3). COs and national partners are jointly responsible for monitoring the country context including early warning monitoring, monitoring the situation of women and children, and monitoring and evaluating the Country Programme. In addition, the CO has direct responsibility for monitoring its own performance. This is generally done through monitoring the quality of programme management, through field visits, Annual and Mid-Term Management Reviews and self-assessment exercises.

Monitoring and Evaluation in Emergencies

28. Accountabilities for M&E do not change in unstable situations or emergencies. The use of an adequately resourced, high quality Integrated M&E Plan, remains key.

29. The preparation of the EPRP, and especially the country emergency profile relies heavily on detailed statistical or qualitative information in areas most prone to disaster. A regularly updated baseline ensures that reliable estimates can be made about the affected population and their
possible needs during the first hours of a disaster. Possible monitoring and evaluation challenges and solutions during a possible emergency need to be reflected in the Emergency Preparedness and Response Plan (EPRP). Ideally, the development of M&E competencies for emergencies should be built into emergency preparedness exercises.

30. The CCCs include numerous accountabilities related to monitoring and evaluation, including for rapid assessments, which provide baseline data for the situation post emergency. Initial or later in-depth assessments may generate additional demands for data collection and analysis. Where indicated, strengthening monitoring capacity should be included in appeals for emergency funding. If required competencies are not present in the office, help from the RO should be requested.

31. Real-Time Evaluations (RTE) take place in the acute phase of an emergency, and are able to influence and direct the wider emergency response. They may focus on operational performance issues as well as the appropriateness of the response, and may look at UNICEF alone or the efforts of multiple partners. An RTE is however not indicated for every crisis, and COs should consult with the Evaluation Office and EMOPS. If agreed, an RTE would take place about 4-6 weeks after the onset of the crisis. Country Offices should plan for an evaluation of all major emergency interventions once the emergency has past.

32. The Emergency Field Handbook (to be released in mid-2005) contains tools for rapid assessments and specific material on other M&E challenges, etc, in many different technical sectors. The Evaluation Training Resource is a set of 7 modules covering all M&E dimensions in regular and unstable situations, and in acute crises. It discusses how to adapt the IMEP and other M&E tools to emergency settings. Monitoring and evaluation are also discussed in training programmes on EPRP and PATH (Principled Approach to Humanitarian Action).

33. For further information, see the “M&E in Emergencies” page on the Intranet. Extensive knowledge about M&E in emergencies is available in other UN organizations (e.g. WFP, UNHCR) and NGOs, who should be consulted as appropriate.
Section 3. Monitoring and Evaluation Responsibilities in UNICEF

34. Monitoring and evaluation activities have been described in Chapters 3 and 4, as they relate to the Country Programme planning and implementation. These included the CCA, the IMEP, the MTRs or Country Programme Evaluation, and the Thematic Evaluation, all at Country Programme level; and programme evaluations and field visits at programme or programme component level. This section describes responsibilities for the planning and management of these monitoring and evaluation activities. Also see E/ICEF/2002/10 on the Evaluation Function in the Context of the Medium-Term Strategic Plan and the subsequent Progress Report on the Evaluation Function in UNICEF E/ICEF/2004/11.

Integrated Monitoring, Evaluation and Research Plan (IMEP)

35. The IMEP is the central tool that helps UNICEF Country Offices and national partners to manage their M&E responsibilities, as established in the CPAP. The IMEP is comprised of two components -- the multi-year IMEP which is prepared and submitted with the CPAP (see Chapter 3, especially Table 3.2.), and the annual IMEP which is prepared with the AMP (see Chapter 4). Both are mandatory and are described in more detail in Chapter 6, Section 6, and the Evaluation Technical Notes.

36. Both IMEPs are essentially calendars of the major research, monitoring and evaluation activities. The multi-year IMEP builds up from the Results Framework and programme logframes and the annual IMEP builds up from as well as feeds into the Annual Work Plans and Annual Management Plan. Both IMEPs must also be consistent with the UNDAF M&E Plan. The multi-year and annual IMEP are of course linked, as are the CPAP and AWP. The annual IMEP gives more detail to those activities identified in the relevant year of the multi-year IMEP and sometimes brings adjustments and additions to major M&E activities. Both IMEPs represent a final stage in the respective programme planning processes to ensure that research, monitoring and evaluation activities are:

- **prioritised** to focus on decision-makers’ most critical information needs, especially given scarce resources;
- **integrated** across programmes and sectors wherever feasible to reduce costs;
- **planned realistically** in terms of timing of activities given end use, practical implementation requirements and capacities of the CO and partners.

37. It is the CO responsibility to ensure that multi-year and annual IMEPs are developed in such a way as to achieve the above results and provide the core information outlined in Chapter 6, Section 6. This requires a good results-based approach to programme planning.

38. Once completed, the multi-year and annual IMEPs serve as management tools that trace out how and when the CO and partners will get the critical information needed for results-based management. The CO is responsible to monitor implementation of and adjust/refine the IMEPs in mid-year and Annual and Mid-Term Reviews.
39. As mentioned in Chapter 3, it is recommended to limit the number of major data gathering activities to no more than 3-5 per year, depending on the CO and partners’ capacities, including financial and human resources. Further, the CO should not plan more research, monitoring and evaluation activities than they were able to implement in the previous year, unless CO capacities have changed dramatically or the previous year was intentionally and strategically a low implementation year for M&E.

**Quality standards**

40. The Representative is responsible for the quality of UNICEF-supported research, monitoring and evaluations. Where necessary, technical support from the regional level, UNICEF HQ, or external sources may be sought.

41. Consistent with UNICEF commitment to a human rights based approach to programming, the organisation promotes a wide participation of stakeholders and especially primary stakeholders in M&E, wherever possible. Wide stakeholder participation is increasingly recognised as being a critical factor in use of M&E conclusions, recommendations and lessons. At the same time, efforts to increase participation must be coupled with attention to mechanisms to ensure the protection of people involved, whether as participants or subjects of M&E activities. There are special implications in this regard for the protection of children participating in monitoring and evaluation. COs are expected to use the [Evaluation Technical Notes, No. 1](#) as a guide in this matter.

42. For evaluation, a number of explicit additional quality standards are well established. UNICEF promotes a utilisation-focused approach to evaluation. When designing, managing or participating in evaluative activities, the CO should consider how each aspect - scope, model, process, methods - will affect use by the intended audience. Consistent with this, the CO and RO have important responsibilities in respect to dissemination, which are discussed below.

43. UNICEF also promotes the use of standard OECD-DAC evaluation criteria including the newer criteria for evaluation of humanitarian assistance (see paragraphs 15 and 16). Depending on timing and purpose of the evaluation, as well as resources available, the scope of the evaluation will be defined, focusing on some criteria and not necessarily covering all. Nonetheless, for all evaluations, COs and partners should explicitly consider each of the standard criteria in articulating the evaluation scope and limitations. Further, all UNICEF-supported evaluations should look at the relevance of programmes/projects in terms of Human Rights-based Approach to Programming and Results-based Management. These are two key strategies in the MTSP and have increasingly wide acceptance in the international community as necessary characteristics of good programming.

44. For all evaluations, COs should use the [Programme Evaluation Standards](#) as a reference in terms of the desired evaluation process and product. These standards are increasingly adopted by national and regional professional evaluation associations. They include standards related to utility, feasibility, propriety and accuracy. COs are expected to use any nationally or regionally-specific variation where they exist and otherwise should use the versions adopted by the African or American Evaluation Associations.
45. Drawing from the Programme Evaluation Standards, UNICEF has developed a set of Evaluation Report Standards which detail the key content and quality aspects of reports. These are useful in clarifying expectations with evaluation teams and should be used by COs in reviewing, accepting or rejecting final evaluation reports submitted. These standards are used by the Evaluation Office in determining which evaluation reports will be included in the Evaluation and Research Database, mentioned below.

46. COs should use the Evaluation Terms of Reference as a means of clearly establishing all of the above quality standards. All evaluation TORs should include the following references to standards:

- in defining the scope and limitations of the evaluation, clear reference should be made to those standard OECD-DAC evaluation criteria that are and are not addressed, and to what degree;
- reference should also be made to the Programme Evaluation Standards as the characteristics of the desired evaluation process and product;
- in defining the evaluation methodology, where any involvement of children is contemplated, Evaluation Technical Note, Issue No. 1 should be mentioned as a required reference;
- in describing evaluation deliverables, clear reference should be made to UNICEF Evaluation Report Standards.

Copies of the above references should be systematically provided to all external evaluators.

47. Finally, COs should establish a mechanism for the quality assurance of Evaluation Terms of Reference. Such a mechanism could entail clearance by the Senior Programme or M&E Officer, a committee of programme staff or could draw on the expertise of external evaluation professionals. The evaluation TORs are the critical management tool for ensuring quality process and product. They provide the means of establishing the quality standards as mentioned above. Further, COs can use TORs to establish milestones and intermediate products throughout the process as entry points to monitor and strengthen quality. Evaluation Technical Notes, No. 2 provides guidance on Evaluation TORs and can serve as a reference for quality assurance.

**Management of monitoring and evaluation resources**

48. The Representative should ensure, and Regional Directors should verify that adequate resources are dedicated to monitoring and evaluation activities. The Executive Board recommends that 2 per cent to 5 per cent of country programme expenditure should normally be devoted to evaluative activities each year, in addition to any support budget covering core staff time contribution to such work (ref: CF/PD/PRO/1986-001). Monitoring activities of COs have increased, in particular with UNICEF support to monitoring the situation of women and children through Multiple Indicator Cluster Surveys, assistance to national partners in reporting to the Committee on the Rights of the Child, and attention to early warning systems development. In support of results based management, the Evaluation Office recommends that the originally defined 2 to 5 per cent be spent specifically on performance monitoring and evaluation. Activities covered by this may include: Annual Reviews, the Mid-Term Review, programme/project evaluations, research and studies, and surveys or other data collection for performance monitoring. In addition, further resource allocations should be made to support
monitoring of the situation of women and children, including through Situation Analysis, as a strategic contribution to national knowledge.

Disclosure

49. Consistent with disclosure provisions established with national partners in the BCA, progress reports and the findings of evaluations of UNICEF-assisted programmes are to be made available to the public. The Mid-Term Review report or CP evaluation, the Annual Review report, as well as programme, project and thematic evaluations are all considered public documents. The evaluative reports which are internal to UNICEF are: the CO Annual Report, the Regional Analysis Report, and reports resulting from the Mid-Term Management Review and the Annual Management Review.

Management of effective learning

50. The Representative is responsible for ensuring that evaluation activities contribute to effective learning, for the CO itself, for programme partners and in conjunction with the Regional and HQ levels, for UNICEF globally. An important aspect in this is the management of the evaluation process, for which existing Programme Evaluation Standards are an important guide.

51. The Country Office also has primary responsibility for disseminating evaluation reports, and especially findings, recommendations and lessons within UNICEF and to programme partners, representatives of primary stakeholders involved in the evaluation, participating agencies and donors. The Regional Office is similarly responsible for promoting and disseminating multi-country evaluations, studies and research, and for reporting on MTRs and major evaluations to the UNICEF Executive Board. Findings can be disseminated through various mechanisms, including formal presentations with national stakeholders, local level community meetings, regional knowledge networks, the CO Annual Report and the Regional Analysis Report.

52. All completed evaluations and studies must be submitted in electronic version to the Regional Director and the Director of the Evaluation Office. (See the Evaluation Report Submission Website on the Intranet.) The Evaluation Offices maintains an Evaluation and Research Database on the Intranet containing summaries of purpose/objectives, methodology, findings/ conclusions, recommendations and lessons learned of all evaluations and studies that meet the above-mentioned Evaluation Report Standards. In addition, the Evaluation Office maintains a repository of all evaluations and studies commissioned by COs, ROs or Headquarters regardless of quality, for accountability purposes.

53. The Representative is also responsible for follow-up to evaluation recommendations, and for seeing that lessons feed systematically into planning processes. This includes specifically a responsibility to ensure that results of all evaluations are discussed and follow-up actions recorded in a meeting of the CMT or other appropriate body, depending on the units, offices or institutions implicated in recommendations. Lessons are also systematically analysed in the Annual Reviews and the Mid-Term Reviews. A distillation of lessons learned must be
summarized in and feed into the formulation of the Strategy Paper and the CPD, the CPAP, and the AWPs.

54. Finally, COs, with regional level assistance where necessary, are responsible for contributing to strengthening national monitoring and evaluation capacities. National monitoring and data collection systems are key elements of national capacity to promote and protect children's rights.
PQAA Checklist – Monitoring and Evaluation

- Have the RO and RMT established strategic evaluation and research themes within the context of the MTSP?
- Has an IMEP been prepared as part of the CPAP and in the context of the UNDAF M&E Plan, and is it being updated annually? Has the RO reviewed and commented on the IMEP?
- What is the proportion of monitoring, evaluation and research activities scheduled for the year that has actually been completed?
- Have at least two major, externally facilitated programme/project evaluation been completed, or are scheduled to be completed before the end of the programme cycle?
- Do all major evaluations involve key stakeholders in design and analysis, and rely on triangulation of data sources and findings?
- Do all major evaluations involve consultation with primary stakeholders, or other forms of active participation, wherever possible? Are mechanisms in place to ensure the protection of those participating in evaluation?
- Is there a mechanism for quality control on the design of major evaluations within the Country Office?
- Has the RO reviewed the design of major evaluations, and offered technical assistance as required?
- Is the MTR supported by formal evaluations?
- Has 2 - 5 percent of country programme expenditure been spent on performance monitoring and evaluation?
- Does the office have a mechanism for reviewing, taking action and follow-up on the findings and recommendations of evaluations?
- Are field visits routinely scheduled and undertaken with programme partners?
- Does the CMT monitor an agreed set of indicators to measure the quality of programme and operations management?
- Does the RO actively support the strengthening of monitoring, survey and research skills among national partners and UNICEF staff in the region?

References and Recommended Reading – Monitoring and Evaluation

- E/ICEF/2002/10, Report On The Evaluation Function In The Context Of The MTSP
- CF/EXD/2005-004, Charter of Authorities and Responsibilities of the Office of Internal Audit
- CF/EXD/1997-01 Information Sharing on Evaluations and Studies
- CF/PD/PRO/ 1998-07, Guidelines for Annual Reviews and Mid-Term Reviews
- Evaluation Technical Notes
- Monitoring and Evaluation Training Resource
- OECD-DAC guidance.
29. In countries where Country Programmes supported by UN agencies are based on an UNDAF, a mandatory UNDAF Annual Review process should be organized jointly by government and the UN agencies. The principle purpose of the review is to assess progress towards achieving expected results as defined in the UNDAF Results Matrix, and to assess the continued relevance of planned results. The annual review process should link to national review processes wherever possible - such as sector reviews, reviews of PRS, reviews of progress towards MDGs and follow-up to the Millennium Declaration. It also takes place in countries affected by emergencies. The degree of formality and elaborateness of the reviews is best determined by the Government and the UNCT. Elaborate reviews require more preparation and need to be justified in terms of an improved future direction of UN cooperation.

30. The UNDAF Annual Review process follows three steps, explained in more detail in the following paragraphs:

   (i) AWP technical-level reviews;
   (ii) theme group analysis;
   (iii) the UNDAF Annual Review Meeting.

31. The AWP technical-level reviews compare achievements against the planned results, activities, inputs and outputs as described in the AWP, with an analysis of the reasons for success or failure. Based on meeting records, notes for the records, field trip reports, progress reports, donor reports, summaries of reviews and evaluations and updated statistical data and indicators, these technical AWP reviews:

   • Assess progress in the achievement of planned results as described in current year’s AWP;
   • Assess the contribution of each AWP to the CP Key Results and – for UNICEF assisted AWPs – the contributions to the organisational Focus Areas expressed in the MTSP;
   • Identify problems and constraints, and the effect of measures already taken to address those;
   • Identify emerging opportunities to accelerate the achievement of the planned results;
   • Assess the usefulness, actual use and status of cash assistance, supply and logistics inputs and technical assistance to government and other partners;
   • Review the implementation of evaluation and research activities planned for in AWPs and the IMEP;
   • Determine if available funds need to be reallocated within the same programmes;
   • Identify major changes in the programme environment, especially in unstable situations, and the likelihood of crisis;
   • Include the review of Procurement Services operations, where they exist (see Chapter 6, Section 14).
   • Provide agencies and implementing partners with conclusions for the next years AWPs.

32. The technical-level reviews of the AWPs should convened by the government, and normally involve all organisations with significant roles or interests in a specific AWP, including
NGOs/CSOs and donors. AWP review meetings may be held separately in specific districts or municipalities where the programme is focusing, allowing for more detailed discussions and participation. For joint programmes, all involved UN agencies and implementing partners should review the relevant AWPs together. The findings and recommendations of the AWP reviews provide inputs into the subsequent thematic reviews, for the UNICEF Annual Report and for the formal UNDAF Annual Review Meeting.

33. Thematic analysis by existing UN theme groups is the opportunity for agencies to collectively assess convergence of agency contributions and overall progress towards UNDAF outcomes. Based on the individual AWP reviews, the thematic analyses should cover:

- Changes in broad planning assumptions, risks and emerging opportunities;
- Continued relevance of UNDAF and agency CP outcomes to national priorities and broader country context;
- Corresponding adjustments to UNDAF and CP outcomes;
- New opportunities for convergence/synergies across programmes, joint programmes and/or M&E activities;
- Necessary revisions to programme approach, cross-cutting strategies, partnerships, resource allocations and the UNDAF M&E Plan.

Each theme group chair will submit to the UNCT a brief (2-3 pages) summary of its conclusions and recommendations for the relevant UNDAF outcome. The theme group summaries will be used in the UNDAF Annual Review Meeting.

34. The formal UNDAF Annual Review Meeting is the once-a-year opportunity for all agencies and national partners to review the contribution of the UNCT to the achievement of national goals based on the UNDAF Results Matrix. This meeting replaces individual agency formal Annual Review Meetings. The UNCT and government will decide on the meeting scope and modalities. The UNDAF Annual Review Meeting should be convened by the government, and normally involves all organisations with significant roles or interests in the Country Programmes, including NGOs/CSOs, UN agencies, and donors. The Regional Office may participate in some cases, but does not generally do so. The meeting will provide:

- A yearly update of overall progress vis-à-vis the UNDAF Results Matrix;
- Validation of conclusions and recommendations that should feed into the preparation of the next round of AWPs.

The UNDAF Annual Review process will also provide inputs to:

- The RC annual report;
- The individual agency country office annual reports;
- Donor reports;
- The UNDAF Evaluation.

35. Annual UNDAF Reviews are held around November or early December. Preparation for AWP reviews usually begins earlier - with the Government, other counterparts and UNICEF Programme Officers compiling material for the reviews.
36. In countries without an UNDAF, technical-level AWP meetings are held, followed by an Annual Review Meeting, convened by the coordinating government ministry. Other development partners contributing to the same results should also be invited. UN agencies and government may choose to combine their separate annual review meetings into one final Annual Review Meeting.

37. Immediately following the UNDAF Annual Review meeting, the UNCT and the concerned government authority confirm the major findings, conclusions, agreed recommendations and follow-up actions through signature of the minutes, or through an exchange of correspondence. The minutes or similar reports should be shared with the UNICEF Regional Director.
Chapter 4, Section 1. Mid-Term Review (MTR) and Country Programme Evaluations (CPE)

51. The Mid-Term Review is held approximately halfway through the CP cycle, usually towards the end of the second half of the mid-year of the CP (ref. PRO/1998-07). In countries with an UNDAF prepared in 2003 or later, the MTR will be conducted as a part of – or feeding into - an UNDAF Evaluation. The principal purposes of the MTR are to:

- examine how the experiences of the CP, at approximately its mid-point, can be used by national partners to improve policies and programmes for the rights of children and women;
- based on a systematic and in-depth review of progress in relation to original CP objectives and expected results, identify and make provisions for mid-course adjustments in the key elements of the CP design as approved by the Executive Board and agreed in the CPAP;
- assess whether modification in the CP results, strategies, distribution of funds between programmes, the CPAP, or the CPMP are warranted as a result of:
  - changes in the country's environment and the situation of children and women, including the likelihood of emergencies;
  - new insights and experience obtained during the first half of the programme cycle;
  - changes in the programme environment (e.g. expected partner contributions not coming forward; new emerging partnerships; changes in access and logistics);
  - changes in national or UNICEF policies and priorities as expressed, for instance, in the PRS or MTSP, or as emerging from the reporting process on the CRC;
- derive major lessons learned so as to improve the quality of programme implementation;
- indicate how these lessons may be applied to the subsequent CP for children and women.

52. The MTR is typically the most substantial of all review exercises in a Country Programme and should be as rigorous and focused as any evaluation. The design of the MTR depends on the specific purpose and scope of the issues to be reviewed, and should aim for both cost-effectiveness and high quality of findings and analysis. The MTR focuses on questions shaping the overall direction and strategic choices of the CP. The following should be considered:

- Preparation of the MTR should be foreseen and incorporated in the AWPs and IMEP;
- The specific purpose, evaluation objectives and priority questions of the MTR must be clearly defined and agreed among partners;
- Stakeholder involvement is required;
- Comparison of findings across different approaches, types or sources of information, methods of data collection and types of analysis;
- The MTR draws on AWP and other monitoring reports, evaluative activities or completed evaluations. At least some of these should provide rights holders’ perspectives. Where no existing evaluative work provides this perspective, new data collection should be contemplated;
- The review process and final report should clearly distinguish the following:
  - UNICEF performance;
  - the shared performance and achievements of the Country Programme partners;
  - Achievements in terms of the quality of both outcomes/impact and processes.
53. As one option, the methodology of the MTR can take the form of a *Country Programme Evaluation (CPE)*. A CPE is externally facilitated and broad in scope. A CPE focuses on information needed for strategic decision-making and improving overall CP performance. The CPE will not lead to lessons and recommendations at the level of programme components or activities; it will rather focus at the level of the strategic contribution of the UNICEF-assisted CP to the evolving situation of children’s and women’s rights. Expected CP outcomes will be reviewed in relation to the MTSP targets, the national and UNDAF outcomes and the country report to the Committee on the Rights of the Child. The CPE will give attention to questions of relevance:

- at the broadest level, is the CP focusing on the right issues?
- specifically, are the key issues for children’s and women’s rights identified in the CCA and CPD still relevant?
- is the positioning of UNICEF support appropriate, in relation to the support from other partners addressing similar issues?
- is the balance between pilots and tested and proven interventions appropriate?
- Are successes of the CP sustainable and can they be effectively and efficiently taken to scale?
- What can be learned from both successes and failures?

A CPE can look at these issues over the period of one or more than one programme cycle.

54. A CPE can be proposed as a methodology for the MTR by the CO, the RO, RMT or Headquarters. It should involve the RO, RMT and/or Headquarters in design and management through some form of steering group. Examples and draft guidance can be obtained from the Evaluation Office.

55. About six months before the final MTR meeting takes place, the government ministry responsible for cooperation with UNICEF and the UNICEF Representative draw up a work plan for the MTR. This should be shared with the RO for comment before finalization. Where an UNDAF Evaluation is planned, the MTR workplan should form part of the overall UNDAF evaluation workplan. The work plan typically includes the date for the MTR meeting, the preparation meetings, schedule and focus of consultations, any evaluations to be carried out with the dates for submission of the evaluation findings, the dates for the submission and review of task force reports, and the date for the preparation of the final draft report in time for the MTR meeting. A good work plan is the equivalent of a Terms of Reference for those involved in the process.

56. Over the preparatory period in-depth analyses of each of the main programme components that make up the CP are carried out. The process culminates in the final MTR meeting.

57. The review meeting lasts from half a day to two days. It reviews the progress of the CP in relation to the original expected results, the resources mobilized and used (compared to those planned), the main results for children and women achieved to date, constraints encountered, the findings of evaluations, and recommendations for corrective action.
58. Following the MTR meeting, two reports are prepared:
- a full report, prepared in cooperation with the government, using the outline set out in the MTR guidelines (ref. PRO/1998-07);
- an internal summary of the full report, of up to five pages, for the use by the Regional Director, including in his/her reporting on the results of the MTR to the Executive Board, and for posting on the UNICEF Intranet.

59. The MTR process and final meeting may take place as part of, in conjunction with, or as an input to an UNDAF Evaluation or a review with national partners of other international programmes of cooperation. In such cases, however, adequate provisions should be made for analysis of the progress and design of UNICEF cooperation specifically, within the wider framework, in order to ensure accountability to the UNICEF Executive Board and funding partners and continuing effectiveness of the resources that UNICEF deploys.

60. Short duration programmes may conduct a MTR as an extended Annual Review process.

61. During or immediately following the last stages of the MTR, a mid-term management review should be carried out, which can in part replace the AMR of that year. The main aim of this review is to analyse the overall management performance of the CO during the first half of the programme cycle in comparison to the planned outcomes in the CPMP, and identify areas for improvement for the remaining part of the programme. In addition to the issues addressed in the AMR, set out above, the mid-term management review would also re-examine the staffing structure of the office and identify any necessary changes in the CPMP, to be proposed and presented to the next regional PBR. A short report of the management review will be written and distributed among the office staff, the Regional Office and Headquarters. Both the review and report will provide inputs to the Country Office Annual Report, and form the basis for the following year's AMP.

62. The MTR process, findings and conclusions, recommendations and lessons learned (for both the Programme Review and the Management Review) should be referred to in detail in the CO Annual Report.

63. In the case of COs responding to the sudden on-set of a major humanitarian crisis that results in a significant increase in Other Resources, and in the absence of an evaluation undertaken with other major humanitarian actors, it is recommended that the CO plan for a major review or Country Programme Evaluation, to take place about 12 months after the on-set of the crisis. If there has been a dramatic change in country context, an externally facilitated CPE may be necessary.
**Chapter 4, Section 1. Programme/project evaluations**

64. For the purposes of accountability, the CO should carry out at least two major programme or programme component evaluations of strategic significance during the CP cycle. To ensure objectivity and credibility, the following should be observed (See: Evaluation Technical Notes No.1: What is in a Terms Of Reference):

- Involvement of key stakeholders – national or other partners, CSOs, donors, UN agencies as well as other key actors – in defining the evaluation purpose, objectives and design, preferably through some form of steering committee;
- Consultation with primary stakeholders (as a minimum, and where possible more extensive involvement of the focus populations);
- Engagement of one or more external evaluators, with at least one of these external evaluators assigned a role as overall facilitator of the evaluation process;
- Triangulation, i.e. the systematic comparison of findings across different approaches, types or sources of information, methods of data collection and types of analysis.

65. CO responsibilities in respect of quality standards, disclosure, dissemination, submission of evaluation reports and management of evaluation follow-up are detailed in Chapter 5.

**Chapter 4, Section 1. Thematic evaluations**

66. The purpose of thematic evaluations is to draw lessons that can be generalized beyond the context of a single project or programme. The theme can be based on:

- a strategy – such as community participation, capacity building, advocacy;
- a priority issue – such as decentralization, civil society partnerships, gender mainstreaming, in-country logistics, or management issues;
- a programme objective or area (including areas related to the MDGs and MTSP priorities).

67. Thematic evaluations are often useful at key review events such as the Mid-Term Review. As part of their evaluation responsibilities, Regional Offices (ROs) and Headquarters also carry out thematic evaluations focusing on strategic issues and feeding into regional and global policy development.
Chapter 6, Section 6. Integrated Monitoring and Evaluation Plan (IMEP)

This section provides additional details for completion of both the multi-year and the annual Integrated Monitoring and Evaluation Plan (IMEP). The IMEP for the UNICEF-supported Country Programme necessarily links to the UNDAF M&E Plan which is covered in the CCA/UNDAF Guidelines. For a stand-alone version of all the IMEP-related guidance in the PPPM, see Evaluation Technical Note no. 4.

Multi-year IMEP

The multi-year IMEP is an essential element of the CPAP and should be submitted with this document to ROs and Headquarters.

Criteria for a good multi-year IMEP

A good IMEP:
- Clearly prioritises a limited number of major research, monitoring and evaluation activities per year, not more than were completed successfully in the previous year, unless CO capacities have changed dramatically or the previous year was intentionally and strategically a low implementation year for M&E.
- Integrates data collection activities across sectors and programmes as this is a critical means of reducing M&E costs.
- Provides a handy reference to monitor information flow for results-based management for the CO, partners and donors.
- Provides a reference for annual planning exercises to ensure that major data collection activities feature in Annual Work Plans and individual work plans.
- Is refined and adjusted on an annual basis according to new information on major events using M&E data or on partners’ major data collection activities, any corresponding shifts in data collection priorities, or changes in CO and partners’ capacities.

The format

Table 3.4 Format of a Multi-year Integrated Monitoring and Evaluation Plan (IMEP)

<table>
<thead>
<tr>
<th>Major events/ processes using research, M&amp;E data</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
<th>Year 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surveys, studies (including SITAN related)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monitoring systems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partners’ major data collection activities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M&amp;E capacity building (UNICEF and partners)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publications (optional)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The multi-year IMEP cannot be developed and would be incomplete without some form of programme Logframes as a basis for identifying indicators and means of verification. Where COs have not used a programme Logframe as detailed in this manual, the UNDAF M&E Framework may be used as a reference. The latter is a required part of the UNDAF M&E Plan and details agency specific CP outcomes and outputs, their indicators and sources of verification.

- **Major events/processes using research, M&E data** are identified in the multi-year IMEP as a planning reference, to keep the focus on users when situating the timing of data collection activities. This section of the IMEP should reflect any relevant events/processes identified in the UNDAF M&E Plan, for example, national or international conferences, MDG reporting, preparation of the PRS, as well as joint review exercises under the UNDAF and preparation of the next CCA, the UNDAF and individual UNCT agency CPs. This section may include some events that are not UNICEF-sponsored, but are opportunities for UNICEF to influence decision-making with data and analysis on the situation of children’s and women’s rights. This section would also likely include the formal UNICEF Mid-Term Review or known major donor reviews. This section in particular is developed further and refined in each annual IMEP.

- **Surveys and studies** refer to major UNICEF-supported surveys (e.g. MICS) or research. This section should reflect key activities contributing to an ongoing monitoring of the situation of children and women. Research will be typically oriented to exploring the underlying causes of a problem/issue where information gaps were identified in a SitAn and/or CCA. Surveys may be national or sub-national in scope; they may focus on the general situation of a population or be more focused on assessing programme outcomes. Joint surveys and studies identified in the UNDAF M&E Plan will appear here. Important surveys conducted by partners such as DHS should be shown in the section “Partners’ major data collection activities”.

- **Evaluations** that ultimately figure in the IMEP should be focused on the most important strategic management questions. Major programme evaluations are likely to precede the Mid-Term Review as well as the UNDAF evaluation and development of the new UNDAF and Country Programme. Evaluations of pilot initiatives will usually be situated early in the programme cycle (year 1 or 2). Ideally, opportunities for joint evaluations will have been identified in the UNDAF M&E Plan and will also appear here.

- **Monitoring systems.** This category includes planned activities to strengthen the monitoring systems at national levels or sub-national levels, e.g. Health Information System (HIS), Country Reporting and Information System (CRIS), DevInfo, sentinel surveillance systems (for disease, nutrition, etc), early warning systems and others.

- **Partners’ major data collection activities.** Other organisations/institutions may be able to provide valuable data for situation monitoring and/or research relevant to the CP. Identifying such data collection activities planned by others helps to avoid redundant efforts and build partnership in data collection. Relevant UN partners’ data collection work will already be identified through the process of preparing the UNDAF M&E Plan. It is also important to identify relevant work of national institutions, NGOs and donors.

- **M&E capacity building.** This listing will show scheduled capacity building activities for improved national data collection and research, for example a MICS training preceding the actual MICS survey; training preceding the introduction of DevInfo, or a longer effort to strengthen a national evaluation or statistics offices or national networks etc.
• **Publications.** This is an optional section of the IMEP. Scheduling the publication of UNICEF-supported monitoring, evaluation and research work in the Multi-Year IMEP facilitates better assessment of work load and resources required in carrying M&E activities through to dissemination.

**The process**
The IMEP flows directly from the results frameworks and programme LogFrames developed in the CPAP. It will also be developed with reference to other agency M&E plans and must be consistent with the eventual UNDAF M&E Plan. (As mentioned above, where COs have not developed full LogFrames, they can refer to the UNDAF M&E Framework, part of the UNDAF M&E Plan.)

**Step 1: Setting parameters**
Two types of parameters need to be established:
- Identify the major events/ processes using research, M&E data that are known over the next 5 years.
- Set a realistic number of major research, monitoring and evaluation activities that the CO and partners can undertake in a year. The recommended limit is 3 to 5 major research, monitoring and evaluation activities per year, depending on the CO and partners’ capacities, including human and financial resources. In assessing overall capacities, one of the most critical and often overlooked issues is the necessary human resources capacity -- both skills and time -- to manage M&E activities and ensure process and products meet quality standards. The CO should not plan more research, monitoring and evaluation activities than they were able to implement in the previous year, unless CO capacity has changed dramatically or the previous year was intentionally and strategically a low implementation year for M&E.
- Identify and list partners’ major data collection activities that are relevant to the CP goals and objectives. As mentioned above, these may already be identified in the process of developing the UNDAF M&E Plan and may already have been cited in the programme LogFrames. Possibilities are also sussed out through further dialogue with UN partners and with other key actors – relevant national institutions, independent research bodies, INGOs, NGOs and donors.

**Step 2: Integrating research, M&E activities across programmes**
Review all major research, monitoring and evaluation activities identified as “Means of Verification” in programme LogFrames. These can be plotted in the multi-year IMEP format and assessed in terms of opportunities for convergence:
- Where the type of data collection activity is the same (qualitative studies, household surveys, evaluations);
- Where the scope is the same (geographic region, population group);
- Where the unit of analysis for data collection (households, communities, service points) is the same, could be the same if the indicators were adjusted, or could feasibly be collected in a same exercise;
- Where the timing is the same or could be adjusted.
This process should produce a reduced list of data collection activities which translates in reduced costs for all partners involved.

Integration of research, monitoring and evaluation activities may come not only through opportunities for convergence across programmes within the CP. This step may also identify new opportunities for joint activities, i.e. integrating UNICEF-supported CP-related M&E activities with research, monitoring and evaluation activities of other partners.

**Step 3: Identify M&E capacity building needs**

For all of the research, monitoring and evaluation activities identified, critical capacity gaps are identified, especially skills/knowledge but also other institutional capacity gaps. Without developing specific activities, the magnitude of effort, resources required and timing of capacity building efforts is traced out and reflected in the multi-year calendar.

**Step 4: Prioritise data collection activities.**

Where the number of research, monitoring and evaluation activities per year exceeds the parameters set from a practical management perspective, priorities are established and lower priority activities dropped. The relative workload for different M&E activities is considered, including options for reducing workload. Prioritisation includes consideration of:

- Strategic importance of data provided – who are the users, how important are the decisions that will be taken;
- Adequacy (from a users’ perspective) of alternative indicators available from existing monitoring systems which can justify dropping more resource intensive M&E activities;
- In light of the above, cost-effectiveness of originally identified research, monitoring and evaluation activities.

**Step 5: Adjusting Logframes and budgeting M&E costs**

Where research, monitoring and evaluation activities in initial programme Logframes have been adjusted or dropped in developing the IMEP, Logframes are revised. The IMEP thus leads to simplification and streamlining of the indicators and means of verification for each programme based on feasibility and priority data and analysis needs. The IMEP also provides a reference to work out at least rough resource provisions for M&E in the CPAP, Summary Budget Table.

Where the process of developing the IMEP points to any major changes in UNICEF and partners contribution to M&E activities already established in the UNDAF M&E Plan, this must be discussed with other partners in the UNDAF. The UNDAF M&E Plan will often be adjusted after agencies have developed their M&E plans for the CPAP. Even where the IMEP leads to no major changes, it is shared with other partners in the UNDAF for general purposes of coordination and information sharing.
The annual IMEP

The annual IMEP is an essential element of the AMP and is a complement to the listing of quality assurance indicators for programme and operations management therein. One without the other is incomplete. The annual IMEP brings together the major activities for M&E of the CP and M&E for UNICEF programme and operations management, laying out a realistic and manageable means of undertaking both.

Criteria for a good annual IMEP

A good annual IMEP:

- Clearly prioritises a limited number of major research, monitoring and evaluation activities for the year, not more than were completed successfully in the previous year, unless CO capacities have changed dramatically or the previous year was intentionally and strategically a low implementation year for M&E.
- Integrates data collection activities across sectors and programmes as this is critical means of reducing costs for M&E.
- Provides a handy concise reference to monitor information flow for results-based management for the CO. It helps the CO staff to see which M&E activities are most needed for which purpose and thus to further prioritise when implementation is slowed or resources must be reallocated. The segments of the annual IMEP related to M&E of the CP are also a reference to be shared with partners and donors and can easily be reproduced eliminating references to UNICEF internal M&E.

Format

Table 6.6.1 Standard format of a annual Integrated Monitoring and Evaluation Plan (IMEP)

<table>
<thead>
<tr>
<th></th>
<th>1st quarter</th>
<th>2nd quarter</th>
<th>3rd quarter</th>
<th>4th quarter</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>External events/ processes using research, M&amp;E data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Internal PP milestones</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Surveys, studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(including SITAN related)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Evaluations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CP monitoring systems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CO internal monitoring systems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Partners’ major data collection activities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M&amp;E capacity building</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(UNICEF and partners)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Publications (optional)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The above is the required format for the annual IMEP. Most of the rows are the same as in the multi-year IMEP. Differences are as follows:
• **Internal Programme Process (PP) milestones** are separated out from External events/processes using research, M&E data. The internal milestones should include: mid-year reviews, stages of the annual review and planning process.

• **CP monitoring systems** are distinguished from monitoring systems focusing on UNICEF CO performance alone. CP monitoring systems will include specific activities to support national or sub-national information systems as identified in the multi-year IMEP.

• **CO monitoring systems** include specific activities to improve or develop internal monitoring systems to track data for CO programme or operation management quality assurance indicators, as identified in the AMP.

• **M&E capacity building** will at this stage be much more clearly defined in terms of what interventions and activities, for what audiences, to build what elements of capacity as well as timing.

With smaller Country Programmes, not every row will ultimately have activities each year. It is however necessary every year to flesh out those rows that serve as planning references – External events/processes using research, M&E data; and Internal PP milestones – as well as Partners’ major data collection activities. For each of these, new information is usually available for the current year and may point to changes in what is planned. In the annual IMEP, activities should be detailed with specific months where relevant.

<table>
<thead>
<tr>
<th>Key results</th>
<th>Indicator, baseline &amp; target</th>
<th>Means of verification (MOV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programme management</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-cutting/process issues</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operations management</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is required as part of the Annual Management Plan to have a listing of programme and operations management indicators. The above format is an optional worksheet that can accompany the annual IMEP, to help specify how each indicator will be monitored. Neither the required listing nor this optional table should be longer than a page for the whole CP for the year.

• **Key results** are included to help COs define what they are trying to achieve before moving to defining indicators. Key results for programme management will most often correspond to the most critical activity outputs lifted from Annual Work Plans. (Considerations in the selection are explained below under Process). COs may also choose to identify a few critical cross-cutting or process results that are pivotal to the quality of the programme – for example, strengthening of gender analysis in assessment and analysis elements in all programmes, integration of protection considerations in all
programme delivery. Finally, the CO will choose a few pivotal outputs in the area of Operations that are critical to achieving programme results.

- **Indicators, Baselines and Targets.** The indicator is the objective measure (e.g. girls enrolment rate), the baseline is the initial level measured for that indicator, and the target is the explicit statement of desired results for the indicator over a specified period of time. Wherever possible the CO should look to simplify, choosing the most cost-effective indicators, which entails thinking ahead to means of verification.

- **Means of verification** (MOVs) for the selected management indicators will most often be simple work plan monitoring or extracting data from existing information systems (e.g. PROMS). Management indicators for cross-cutting issues are often covered through field monitoring or programme/project reporting mechanisms. A few indicators may require some additional form of data collection for monitoring. It is important that additional monitoring activities are not adopted without weighing the cost-effectiveness of the information provided for decision-making.

**The process**
The development of the annual IMEP is integral to the development of AWPs and the AMP.

**Step 1: Refining the multi-year IMEP**
Prior to working on AWPs and the AMP, it is important to review the multi-year IMEP to check on its continued relevance. New events or processes that draw on M&E data, i.e. where UNICEF has an opportunity to influence decision-making with good data on the situation of children and women, may be identified. This may add to the demands for results from research, M&E activities. Additional major data collection activities of partners’ may also be identified, potentially reducing M&E demands on UNICEF and partners.

**Step 2: Integrating M&E in AWPs**
Each programme manager will refer to the multi-year IMEP, to ensure that relevant research, monitoring or evaluation activities already identified are integrated in the Annual Work Plans (AWP) as discrete activities with budgeted resources. Where one data-collection activity cuts across programmes, a decision is taken as to who will manage and which AWP budget will cover the activity.

**Step 3: Defining programme and operations management indicators for the AMP.**
In developing the AMP, the CO develops a set of quality assurance indicators for programme and operations management. These can be worked out using the format in Table 6.6.2 above. In refining a manageable list of indicators it is useful to focus on:

---

**Work planning for M&E activities**

In planning for major surveys, research and evaluation activities, it is useful to consider the time and major resources required for each of the following common tasks: developing the TOR, selecting the data collection team, further methodology design and testing, data collection, data analysis, dissemination workshop, publication, etc. For more guidance see *UNICEF M&E Training Resource*, Module 3, Managing M & E Activities.
• Results that are indicative of the overall progress (e.g. no. of children immunized/month as a programme result; timeliness of cash and supply deliveries as an operations management result);
• Results that are necessary conditions for other outputs to be realized this year or next (e.g. development of a new curriculum which kicks off a series of roll-out activities in education);
• Result areas that are considered problematic, where there are known challenges in delivering an input or implementing an activity (e.g. development of a human rights network in a politically charged context as a programme result; increasing attention to marginalized populations as a cross Cutting result; or reducing outstanding Cash Assistance to Government as a management result).

Note that in unstable or crisis contexts, COs will often identify a few contextual assumptions pivotal to programme implementation – e.g. coverage of access -- that will similarly be reviewed periodically to check its affect on programme progress and adjust operational strategies as necessary.

Step 4: Integrating the annual IMEP
With AWPs developed and a list of key indicators and MOVs identified, the CO must then assemble all major research, monitoring and evaluation activities into the annual IMEP format. This will include those originally identified in the multi-year IMEP as well as any new ones identified through the AWP and AMP processes. As with the multi-year IMEP, opportunities for integrating M&E activities across sectors/programmes are examined. Similarly, where the major research, M&E activities planned exceeds what is considered manageable, the CO must prioritise and eliminate lower priority activities.
United Nations Children’s Fund
Executive Board
Annual session 2002
3-7 June 2002
Item 5 of the provisional agenda*

Report on the evaluation function in the context of the medium-term strategic plan**

Summary


Following the introduction, chapter II provides the background to the report. An overview of the evaluation system in UNICEF and the accountability framework for evaluation are presented in chapter III. Recent measures taken to strengthen the evaluation function are described in chapter IV. The proposal for a multi-year evaluation plan in support of the MTSP is presented in chapter V. Chapter VI contains a draft recommendation for Executive Board approval.

** The need for extensive consultation within the secretariat delayed the submission of the present report.
# Contents

<table>
<thead>
<tr>
<th>Paragraphs</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Introduction</td>
<td>1–2</td>
</tr>
<tr>
<td>II. Overview</td>
<td>3–11</td>
</tr>
<tr>
<td>A. Background</td>
<td>3–6</td>
</tr>
<tr>
<td>B. Evaluation in the context of the medium-term strategic plan</td>
<td>7–11</td>
</tr>
<tr>
<td>III. UNICEF evaluation system</td>
<td>12–34</td>
</tr>
<tr>
<td>A. Evaluation within the performance monitoring and oversight framework of UNICEF</td>
<td>12–17</td>
</tr>
<tr>
<td>B. Purpose of the evaluation function</td>
<td>18–22</td>
</tr>
<tr>
<td>C. Findings from the peer review</td>
<td>23–24</td>
</tr>
<tr>
<td>D. Stratification of the evaluation system</td>
<td>25–29</td>
</tr>
<tr>
<td>E. Accountability for the evaluation function</td>
<td>30–34</td>
</tr>
<tr>
<td>IV. Measures taken to strengthen the evaluation function</td>
<td>35–49</td>
</tr>
<tr>
<td>A. Weaknesses that need to be addressed</td>
<td>35–39</td>
</tr>
<tr>
<td>B. Strengthening of in-country evaluation capacity</td>
<td>40–41</td>
</tr>
<tr>
<td>C. Strengthening of the country offices</td>
<td>42–43</td>
</tr>
<tr>
<td>D. Strengthening of the regional offices</td>
<td>44–45</td>
</tr>
<tr>
<td>E. Strengthening of New York headquarters</td>
<td>46–47</td>
</tr>
<tr>
<td>F. Fortifying management of the evaluation function</td>
<td>48–49</td>
</tr>
<tr>
<td>V. Multi-year evaluation plan in support of the medium-term strategic plan</td>
<td>50–57</td>
</tr>
<tr>
<td>A. Evaluation of the organizational priorities</td>
<td>52–53</td>
</tr>
<tr>
<td>B. Evaluation of the country programme of cooperation</td>
<td>54</td>
</tr>
<tr>
<td>C. Evaluation of organizational performance</td>
<td>55</td>
</tr>
<tr>
<td>D. Easier access to the organizational memory</td>
<td>56–57</td>
</tr>
<tr>
<td>VI. Draft recommendation</td>
<td>58–59</td>
</tr>
</tbody>
</table>
I. Introduction


2. The Executive Board last considered a report on the evaluation function, entitled "Overall progress in the implementation of evaluation activities in UNICEF" (E/ICEF/1992/L.9), at its 1992 regular session (E/ICEF/1992/14, decision 1992/24). In response to Executive Board decision 1995/8 (E/ICEF/1995/9/Rev.1), the secretariat submits annually to the Board at its annual session a summary of the outcome of mid-term reviews (MTRs) and major evaluations of country programmes, specifying, inter alia, the results achieved, lessons learned and the need for any adjustments in the country programmes. In addition, the Executive Director reports to the Executive Board on evaluation matters in part II of her annual report. In 1999, the Executive Board decided that starting from 2000, information in part II of the Executive Director's report should be presented in a way that facilitates monitoring of progress in achieving the objectives of the programmes and activities within the framework of the organizational priorities in the medium-term plan (MTP) for the period 1998-2001 (E/ICEF/1998/13 and Corr.1 and E/ICEF/1999/7/Rev.1, decision 1999/7).

II. Overview

A. Background

3. In decision 1992/24, the Executive Board reaffirmed its decision 1990/4 (E/ICEF/1990/13) that a past review of evaluations and their use, as well as a summary of the evaluation plan and structure, be included in all country programmes. In that same decision, the Executive Board also decided the following: that this evaluation plan include evaluations in all programme areas assisted; that in addition to being a project-focused effort, evaluation at the country programme level should increasingly address programme-level activities; that UNICEF should make available an enhanced evaluation database to monitor evaluation implementation and to facilitate the learning process; that the necessary financial and staff resources be available for implementing evaluation plans and for monitoring the use of results; that a three- or four-year rolling evaluation plan be established; that joint evaluations with donors be intensified; and that collaboration on evaluation be strengthened with Governments in order to address the capacity-building and institutional-strengthening requirements through the country programme and that priority in this regard be given to sub-Saharan Africa.

4. Pursuant to Executive Board decisions and recommendations from external auditors’ reports and the multi-donor evaluation of UNICEF (E/ICEF/1993/CRP.7), the Deputy Executive Director, Programmes, announced the formation of the Evaluation and Research Office in his Executive Directive of June 1993 (CF/EXD/1993-006). That decision was taken to better reflect the commitment of UNICEF to strengthening national capacities for essential national research for
children and women. It also reflected measures for strengthening the overall evaluation capacity of UNICEF and improvement of the function in support of programme planning.

5. The Executive Board, during its annual session of June 1998, approved the new organization of UNICEF (E/ICEF/Organization/Rev.3 of 24 April 1998) in the context of the implementation of management excellence as well as of the 1998-1999 biennial support budget. UNICEF headquarters was reorganized to focus on strategic, policy, advocacy and oversight functions. This was done taking into account that UNICEF had always been a decentralized, field-based organization and that headquarters structures worked together to best support and strengthen country programmes and the effective delivery of the UNICEF mission. The Evaluation, Policy and Planning (EPP) Division was created to provide technical leadership in monitoring and evaluating the effectiveness and efficiency of organizational performance in monitoring the global situation of the child; to ensure that the results of evaluations were fed into the development of organizational policies and strategies; to analyse the impact of social and economic trends and policies on children; and to coordinate strategic planning and the development of MTPs for the organization. As a consequence of the reorganization, the Office of Evaluation and Research became a unit of EPP.

6. In December 2001, in the context of the approval of the MTSP (decision 2001/22) and the 2002-2003 biennial support budget (E/ICEF/2001/AB/L.10 and decision 2001/13), the Executive Board endorsed the reorganization of the headquarters Programme Group based on the results achieved and experience gained from the former structures. Responding to the need to use the evaluation function more strategically and to provide technical support to fortify performance assessment, the Evaluation Office was given the status of a separate office with increased resources, reporting to the Deputy Executive Director, Programme and Strategic Planning. This measure also enables UNICEF to be more in conformity with international professional standards regarding the positioning of the Evaluation Office within the organization.

B. Evaluation in the context of the medium-term strategic plan

7. The MTSP combines a reinforced results-based management approach and a human rights-based approach to programming. Building on the lessons learned from implementation of the MTP, the new plan establishes five organizational priorities, more clearly defines objectives and indicators, and strengthens the strategic use of the evaluation function. For the first time, a plan has been proposed for the evaluation of the MTSP.

8. The MTSP indicates that evaluation will focus more on the country programme level and on institutional management of the organization as a whole. It will look at the rationale, effectiveness and administrative efficiency and economy of activities undertaken or supported by UNICEF. Evaluation will support accountability and results-oriented performance.

9. Country programme evaluations will gradually be strengthened. During the first two years of the MTSP, the Evaluation Office will develop basic principles and methodologies and conduct a limited number of field tests, taking into account
previous work on the subject. From the third year of the MTSP, regional offices will gradually assume responsibilities in this regard.

10. A special effort has been made to formulate the MTSP so that organizational priorities express the strategic intents pursued from an institutional perspective and so that indicators serve as benchmarks for the assessment of organizational performance. At the end of the third year of the four-year period, a review of the implementation of the MTSP will assess progress made towards the organizational priorities. MTRs and major evaluations of country programmes will inform this review. Lessons learned from the review will be used for the development of the next MTSP.

11. The evaluation plan for the duration of the MTSP will cover key themes and topics of strategic significance. The organizational priorities of the MTSP will guide the selection of thematic evaluations to be undertaken at country, regional and global levels. Such evaluations will be conducted with an emphasis on programmes, strategies and policies. Topical evaluations will address a variety of cross-cutting themes as well as UNICEF organizational effectiveness. Implementation of the evaluation plan will, in some cases, involve partnerships with other United Nations agencies and/or governmental and non-governmental organizations. Findings will be stored in an on-line electronic database, and learning workshops will be part of the dissemination of evaluation results.

III. UNICEF evaluation system

A. Evaluation within the performance monitoring and oversight framework of UNICEF

12. During the third regular session of 1997, the Executive Board endorsed the framework of roles, responsibilities and accountabilities for performance monitoring and oversight (E/ICEF/1997/AB/L.12 and E/ICEF/1997/12/Rev.1, decision 1997/28). Performance monitoring and oversight were major themes throughout the management excellence process in UNICEF. Their purpose is to ensure high quality and responsive programmes through the responsible use of resources for the maximum benefit of children and women.

13. Performance monitoring and oversight feature in all aspects of UNICEF work. The UNICEF system of oversight is a cyclical process involving assessment of programme and operational performance against organizational priorities and objectives generated by the planning process. The answer to the question “How are we performing against what we set out to achieve?” is obtained through “performance monitoring”, a management function carried out in offices throughout UNICEF, and “oversight”, separate independent mechanisms to assess programme and operational performance.

14. The fulfilment of accountabilities within UNICEF is assessed through a dual system of performance monitoring and oversight. Performance monitoring includes all tasks associated with supervision. It is a management function assigned at all levels of the organization. Oversight of these management functions is maintained through independent internal audit and investigative functions carried out within UNICEF, and by mandated external bodies within the United Nations system.
Implementation of accepted recommendations from oversight activities is then, in turn, a responsibility of line management.

15. The evaluation function in UNICEF is both a mechanism for providing oversight at country, regional and headquarters locations and an instrument that allows organizational learning through the identification of lessons and good practices. Evaluations are conducted as a component of performance monitoring to assess whether UNICEF programmes achieve their objectives and are effective and relevant, and to distil lessons for improved programming, strategic planning and policy development. Evaluations are also commissioned by the Evaluation Office as a component of the independent oversight activities of UNICEF.

16. The research function also contributes to organizational learning and knowledge acquisition. It enhances effectiveness during the design of approaches, policies, strategies and programmes. Research is concerned with testing and understanding basic models and approaches, and is based on scientific methodologies. In UNICEF, the Innocenti Research Centre, Programme Division, the Division of Policy and Planning and country offices conduct research studies and contribute to organizational learning.

17. Thus, the evaluation function is one of many functions within the performance monitoring and oversight system. Evaluation is not an inspection, nor is it an audit. It should not be confused with monitoring, which is a management function of self-assessment and reporting. Evaluation should not be expected to yield scientific findings such as those emanating from fundamental research.

B. Purpose of the evaluation function

18. In the Secretary-General’s bulletin on the regulations governing the methods of evaluation (ST/SGB/2000/8) issued on 19 April 2000, pursuant to General Assembly resolution 54/236 of 23 December 1999 and its decision 54/74 of 7 April 2000, the objectives of evaluation are defined in regulation 7.1:

(a) To determine as systematically and objectively as possible the relevance, efficiency, effectiveness and impact of the Organization’s activities in relation to their objectives;

(b) To enable the Secretariat and Member States to engage in systematic reflection, with a view to increasing the effectiveness of the main programmes of the Organization by altering their content and, if necessary, reviewing their objectives.

19. The report on the “Implementation of management excellence in UNICEF” stated that “the evaluation function in UNICEF is both a mechanism for providing oversight at country, regional and headquarters locations and an instrument that allows organizational learning through the identification of lessons and good practices” (E/ICEF/1997/AB/L.12, paragraph 4).

20. Hence, the evaluation function has many purposes. Evaluation is essentially about identifying and understanding results and their impacts, aiming at the provision of useful information and best alternatives to inform decision-making. Its intent is to enable learning-by-doing, thus improving results-oriented activities by re-engineering ongoing activities or improving the design of new ones. The
formative evaluation process is participatory and is an empowerment tool fostering fairness and impartiality, enlarging the potential for consensus-building. Finally, evaluation is about accountability because it focuses on results achieved or not achieved and on explaining what has been achieved and why. It shows what decisions/actions were taken in light of what happened. Most of all, it enables the provision of information on results and learning to stakeholders and the public.

21. In summary, evaluation is the function that examines a policy, a strategy, a programme or an activity/project by asking the following questions: Are we doing the right thing? Are we doing it right? Are there better ways of doing it? It answers the first question by proceeding with a reality check, by examining the rationale or justification, and by assessing relevance in relationship to the fulfilment of rights. The second question is answered by examining effectiveness through the lenses of the pertinence of the results achieved and by assessing efficiency with the review of the optimization of the use of resources. The third question is dealt with by identifying and comparing alternatives, by seeking best practices and by providing relevant lessons learned.

22. Professional experience and learning point to the following six key characteristics for good evaluations:

(a) Impartiality: neutrality and transparency of the evaluation process, analysis and reporting;
(b) Credibility: professional expertise, methodological rigour, participation and transparency;
(c) Usefulness: timeliness for decision-making, and clear and concise presentation of relevant facts;
(d) Participation: reflection of different interests, needs and perceptions, and sharing among stakeholders;
(e) Feedback: systematic dissemination of findings to stakeholders and use in decision-making;
(f) Value-for-money: value-added outweighs the costs.

C. Findings from the peer review

23. In December 2000, a peer review was conducted of the evaluation function in UNICEF. The heads of evaluation of the United Nations Development Programme, the United Nations Population Fund, the World Food Programme, the Office of the United Nations High Commissioner for Refugees and the World Bank, as well of the Director of the Office of Internal Audit, proceeded with a comparative examination of the evaluation function. The review concluded with the following findings:

(a) There is a lack of a common set of norms and standards that govern evaluation functions within the United Nations system in spite of the General Assembly resolution requesting harmonization;
(b) The introduction of results-based methodologies has significant implications, and the traditional oversight approaches need to be reassessed;
(c) Country programme evaluations need to be recognized as a unit of evaluation;
(d) The issue of attribution needs to be revisited in the context of partnership approaches;
(e) The role and level of central evaluation offices respond to different organizational expectations within the United Nations system; some are independent, while others are twinned with audit or other oversight functions;
(f) Most evaluation units within the United Nations system are more centralized and many are oriented to policy-making, whereas evaluation in UNICEF has been oriented towards programme guidance.

24. The peer review also referred to the principles for evaluation of development assistance issued in 1991 (OECD/GD(91)208) and reassessed in 1998 by the Development Assistance Committee of the Organisation for Economic Co-operation and Development (DAC/OECD). These principles reveal a strong consensus among the heads of evaluation of the bilateral agencies on the following principles:
(a) Agencies should have an evaluation policy with clearly established guidelines and methods, and with a clear definition of its role and responsibilities and its place in the organizational structure;
(b) The evaluation process should be impartial and independent from the process concerned with policy-making and the delivery and management of development assistance;
(c) The evaluation process must be as open as possible, with the results made widely available;
(d) For evaluations to be useful, they must be used; feedback to both policy makers and operational staff is essential;
(e) Partnership with recipients and donor cooperation in evaluation are both essential; they are an important aspect of in-country institutional-building and coordination, and may reduce administrative burdens on countries;
(f) Evaluation and its requirements must be an integral part of planning from the start; clear identification of the objectives that an activity is to achieve is an essential prerequisite for any evaluation.

D. Stratification of the evaluation system

25. In UNICEF, there are three levels where results are being achieved. They are: the local activity or project level; the country programme of cooperation level; and the organizational management level, including the organization’s own organizational performance. These levels correspond well with the accountability framework reflected in the organization of UNICEF (E/ICEF/Organization/Rev.3). For each level, there is a management cycle consisting of the five phases of planning, programming, implementation, monitoring and evaluation.

26. At the activity/project level, a diagnosis of the need is made and an expected result answering to the identified need is articulated as the objective of the project/activity, together with performance indicators and risk assumptions. This is
the planning phase that is completed in tandem with the programming phase. The latter consists of the preparation of an explicit work breakdown structure, a schedule of events, a budget and a matrix of accountability related to the undertaking of each task, as well as the overall management of the activity/project. Implementation is carried out by the programme partners, contractors, or directly by UNICEF staff. Monitoring ensures the measurement of progress and reports the gaps, enabling the orientation of activity/project implementation according to the plan or the realignment of the activity in order to maximize impact and optimize the use of resources. At this level, evaluation is used, in a participatory fashion, to examine results, the relevance of the activity/project design in light of the needs, the effectiveness and sustainability of the effects, the efficiency of management, and economy in the use of resources for the purpose of informing decision-making and learning.

27. The results management framework at the level of the country programme of cooperation also entails the same five management phases. During the planning phase, a situation analysis is conducted, the rights-based approach reveals the gaps and areas of priority, alternative interventions are considered, and a programme proposal is structured and submitted to the Executive Board for approval. During the programming phase, an integrated monitoring and evaluation plan (IMEP) is prepared. The IMEP process strengthens the rights-based and results-oriented focus of the master plan of operations. The IMEP makes explicit the objectives tree of the country programme; identifies the key performance indicators and risks; and provides a systematic approach to monitoring, evaluation and research in support of programme management. Implementation is monitored by means of annual country programme reports and periodic audits. The regional directors report annually to the Executive Board on MTRs and major evaluations of country programmes. Formal comprehensive evaluations of country programmes of cooperation now being piloted are expected to be conducted more systematically in the future.

28. At the level of the organizational management of UNICEF activities, the same five management phases are being put in place with more rigour. The MTSP is the business plan of the institutional priorities of UNICEF. It is based on a diagnosis emanating from the end-decade review and the global needs expressed by member countries in international forums that have led to the setting of global targets such as the Millennium Development Goals. The multi-year funding framework integrates the major areas of action, resources, budget and outcomes, in compliance with Executive Board decision 1999/8. Annual reports submitted by the Executive Director to the Executive Board provide progress reporting on implementation. The organizational performance of UNICEF is assessed by means of the mid-term review of the MTSP and implementation of the multi-year evaluation plan.

29. Thus, there is an evaluation function being performed at each of the three results management levels. The main purpose of the evaluation function is to inform decision-making and distil lessons learned to be used for future planning at each level of results management within the organization. It should be noted that different evaluation approaches and methodologies need to be applied in order to respond to the needs of each level of management. Moreover, for each level, the evaluation function addresses the needs of different networks of decision makers. At the activity/project level, the users of evaluation are the stakeholders, the project team and the country management team (CMT). At the level of the country programme of cooperation, those directly interested in evaluation of the country
programme are the national authorities, the CMT, the regional office and headquarters. Organizational management-level evaluations are of interest to the Executive Board, senior management at headquarters and regional offices.

E. Accountability for the evaluation function

30. The decentralization of the evaluation function is a singular characteristic of the UNICEF evaluation system compared to other international organizations. The country office conducts most of the evaluation work. Regional offices provide oversight and support for evaluations undertaken by the country offices. Regional offices also conduct thematic evaluations related to their regional strategies. Headquarters divisions undertake evaluations relating to their areas of expertise. The Evaluation Office provides functional leadership and overall management of the evaluation system. It also conducts and commissions evaluations.

31. In each country office, an evaluation focal point is accountable to the country representative, who reports annually to the regional director on evaluation findings. Each regional office has a monitoring and evaluation officer who coordinates evaluation work performed by the country offices and their own regional office. The regional director provides annually a report to the Executive Board on MTRs and major evaluations. From a headquarters perspective, the Executive Director reports on evaluation matters to the Executive Board in the context of part II of her annual report.

32. It is the role of UNICEF country representatives to ensure that adequate UNICEF staff resources are dedicated to evaluation, that communication with government officials and other partners facilitates the evaluation process, and that evaluation findings inform the decision-making process. Particularly critical in this is the oversight responsibility that UNICEF representatives have concerning the articulation of the IMEP and the respect for quality in the conduct of evaluations (according to the standards and norms set by the Evaluation Office). The representatives also have to ensure that their annual reports highlight the main evaluation findings and that evaluation reports are registered in the UNICEF evaluation database. Key evaluation activities carried out by the country office are to: develop and update an IMEP; ensure the conduct of evaluations and studies in accordance with the plan, including design, coordination and implementation; ensure the quality and appropriate use of evaluative activities, including MTRs; monitor the effectiveness and relevance of the UNICEF country programme; ensure follow-up of evaluation recommendations; and channel evaluative results into the development of programme strategies and policies.

33. The evaluation function at the regional level focuses on strengthening the monitoring and evaluation capacities of UNICEF offices and their government counterparts through the following: coordination with the Evaluation Office at headquarters; preparation of regional evaluation plans; provision of technical assistance and oversight to support effective monitoring and evaluation of country projects and programmes; and preparation and review of training plans. In accordance with their regional evaluation plans, the regional offices undertake thematic evaluations. They ensure the contribution of their respective region to global evaluations led by the Evaluation Office, and are also responsible for the conduct and oversight of country programme evaluations. The Regional Management Team plays a key role in establishing regional evaluation priorities.
Key evaluation activities carried out by the regional office are to: coordinate the review of MTRs and major evaluation reports in the region, in cooperation with Programme Division and the Evaluation Office, and submit reports on results to the Executive Board; monitor evaluation activities and review evaluation reports in the region to ensure quality and relevance; ensure the evaluation of regional and multi-country initiatives within the region; synthesize evaluation results and lessons within the region; monitor the quality and use of evaluation results to strengthen programmes within the region; and facilitate the exchange of relevant information and experience in the region.

34. At headquarters, the Director of the Evaluation Office is responsible for overall development and implementation of the evaluation work plan, and reports to the Deputy Executive Director, Programme and Strategic Planning. The Evaluation Office has the following accountabilities: to conduct evaluations; and to seek to reinforce the organization’s capacity to address evaluation needs, with an emphasis on the requirements of country offices and capacity-building in countries, in accordance with decisions made by the Executive Board and the Economic and Social Council. The Office provides technical guidance for a comprehensive system of performance management and leadership in the development of the corresponding approaches, methodologies and training for policy, strategic, programme and project evaluations. It monitors and reviews the quality of UNICEF-sponsored evaluations. The Office advises UNICEF senior management on the results of evaluations and related studies, with particular attention to the relevance of these results for organizational processes and policy development. The Office maintains the organizational database of evaluations and research studies, ensures access by UNICEF offices and promotes their dissemination and utilization through all available channels. The Office also collaborates with other United Nations agencies to increase the harmonization of evaluation activities and guidelines through the Inter-agency Working Group on Evaluation. The Evaluation Office is responsible for coordination at the global level with donors, major non-governmental organizations and other partners on the evaluation activities of programmes funded by donors or executed jointly with other organizations.

IV. Measures taken to strengthen the evaluation function

A. Weaknesses that need to be addressed

35. The last systematic and comprehensive review of the quality of evaluations conducted by UNICEF was undertaken in 1995. The objective of that review was to assess the relevance, quality and usefulness of UNICEF-supported evaluations and studies. Other objectives of the review included the estimation of the proportion of impact evaluations and the usefulness of non-impact evaluations and studies, the cost/benefit ratio, the issue of quantitative versus qualitative approaches and the role in capacity-building, and the validation of the evaluation database in terms of the classification of the reports registered.

36. The reviewers concluded that the database was fairly accurate in the classification of the reports. It was found that 15 per cent of all reports registered and 35 per cent of the evaluations recorded dealt with the impact of UNICEF-funded activities. The review showed that 91 per cent of the non-impact evaluations...
and 31 per cent of the studies had relevant findings for possible reformulation of UNICEF-supported projects or programmes. Only 10 per cent of all reports were deemed worthless, and over 27 per cent of the sample reviewed were judged unjustified in terms of costs relative to objectives and actual outcomes. Very few studies and evaluations appeared to have specific and substantial capacity-building components. Six out of every seven studies used quantitative methods, but useful qualitative insights were also derived from most of the reports. Regarding the overall quality of the reports, 3 per cent were inadequate, 29 per cent were poor, 28 per cent were considered fair, 25 per cent were assessed as good and 15 per cent were rated excellent. The reviewers felt that the most common reasons for inadequate reporting might have been the lack of communication between consultants and UNICEF officers, and the lack of foresight (no baseline data, insufficient time and resource allocation or inadequate competence of the investigators in the field under study).

37. In 2000, a review of the UNICEF evaluation database was conducted. It found that the database had recorded some 11,000 evaluations and studies of UNICEF-supported projects and programmes since 1987. In 1992, the Executive Board requested the development of an enhanced database (decision 1992/24). A test version was first released in 1993, under the DOS environment, followed by a complete release in 1994. A CD-ROM was distributed in 1995 containing all of the information in the database. A new version was prepared in 1996 in the Windows format based on inputs from country and regional offices. Updated CD-ROMs were released in 1997, 1998 and 1999. At the beginning of 2002, the Intranet version of the evaluation database was released, allowing real-time, on-line access. Despite the long history of the evaluation database, the 2000 review revealed that it was not as widely known or used in UNICEF as had been expected.

38. In 1990, the Executive Board, noting the importance of evaluation as a management tool in improving programme effectiveness, requested that monitoring and evaluation plans and structures be elaborated and included in all country plans and major projects presented to it (decision 1990/4). In 1993, the Executive Board requested the Executive Director to ensure that country programme evaluations became an integral part of the country programme exercise, with a view to providing better assessments of the performance of the Fund (E/ICEF/1993/14, decision 1993/5). In the 1990s, the Office of Evaluation and Research piloted five evaluations of country programmes. Some country offices also experimented with approaches to the self-evaluation of country programmes. In 2001, the Evaluation Office undertook the evaluation of two country programmes. It is presently conducting the evaluation of the programme of cooperation with the Pacific island countries at the request of the Executive Board.

39. Due to the lack of systematization of the use of the evaluation function at each level of management, evaluations were being conducted mostly at the project level. This explains why over the past years, there has been little reporting on global evaluations. In addition, the lack of systematic use of country programme evaluations explains the discrepancies in the level, depth and scope of the annual MTRs and major evaluations. With the introduction of the MTSP-related multi-year evaluation plan, the eventual conduct of country programme evaluations by regional offices and the increase in the quality of project/activity evaluations led by country offices, there is a high expectation that organizational reporting on results at all levels of management will be enhanced significantly. The challenges during the
MTSP period require that UNICEF go beyond the number, quality and use of evaluations at the individual project level to managing the evaluation process itself more systemically and effectively at country, regional and global levels. More emphasis needs to be placed on assessing the results, impact and effects of programmes and on evaluating country programmes as a whole, as well as assessing the impact of global policies.

B. Strengthening of in-country evaluation capacity

40. Two Economic and Social Council decisions request that particular attention be given to capacity-building in member countries. The first decision states that greater emphasis should be given to helping countries evaluate their programmes themselves and strengthen their own continuing evaluation machinery. The second decision indicates that further work should be undertaken in evaluation, particularly in relation to strengthening national capacities for evaluation and laying the basis for sound programming. UNICEF support to national evaluative activities is anchored at the country level, where the UNICEF country office plans, implements, monitors and follows up on activities of cooperation with the Government.

41. At the regional level, UNICEF has been supporting the formation of evaluation associations, facilitating the collaboration and mutual strengthening of professional evaluators at the national level. In compliance with a decision of the Executive Board requesting that particular support be provided to African countries, the Eastern and Southern Africa Regional Office has been involved in the formation and strengthening of the African Evaluation Association and has provided secretarial support for the articulation of a professional code, the setting up of an evaluators roster and the hosting of annual meetings of the Association. Other regional offices have also been associated with the activities of regional evaluation associations, such as the Central American Evaluation Society and the Australian and Asian Evaluation Association.

C. Strengthening of the country offices

42. At the country programme level, the Evaluation Office has promoted the systematic use of the IMEP within the programme management cycle. Such an evaluation plan is a prerequisite to the gathering of key information necessary for a subsequent evaluation of the country programme. The IMEP is used to strengthen and link planning, monitoring, evaluation and research components of country programmes, and to provide a rational approach to trace relevant information supporting performance-related decision-making. The IMEP has also been adapted as a management tool for global-level programmes and initiatives, in particular for UNICEF efforts on HIV/AIDS. During 2001, IMEP methods and procedures were refined and integrated into the programme process and procedures training manuals. The Evaluation Office is further supporting the generalization of IMEP practices through the facilitation of regional training workshops as well as the dissemination of good practices.

43. The system of evaluation focal points in country offices was initiated in 1987 to strengthen the management of evaluation processes. In each office, a professional staff member is designated as the contact officer for evaluation matters. These focal
points have the following responsibilities: to assist in designing, updating, implementing and monitoring plans to promote and support evaluations; to share evaluation results and disseminate lessons learned within the office and with partners for use in the programming process and project planning; and to prepare proposals and coordinate the training of both government and UNICEF staff for improved monitoring and evaluation. In order to reinforce the identification of skills required, the Evaluation Office is preparing a competency profile for evaluation officers which will be used as technical selection criteria for staffing purposes and also as a benchmark for identifying training requirements.

D. Strengthening of the regional offices

44. The multi-donor evaluation of UNICEF noted that a gap exists in the UNICEF accountability system at the level of accounting for the impacts and effects of UNICEF-supported programmes. Although UNICEF is an agency with complex partnership arrangements and goals, more emphasis must be placed on evaluating country programmes. This emphasis can be enhanced by the development of a clearer and stronger role for headquarters and regional offices in ensuring that evaluation is an integral part of country programme management and in playing a challenge function to ensure that country office staff address strategic-level issues in evaluations. In collaboration with regional offices, the Evaluation Office is conducting pilot evaluations of country programmes. A methodological approach for the conduct of country programme evaluations will be prepared in 2003. It is expected that by 2004, the regional office will gradually assume responsibility for conducting the evaluation of country programmes more systematically.

45. Over the years, regional offices have given attention to the function of monitoring the situation of children and programme performance. There is a need to strengthen the capacity of regional offices in evaluation. Regional monitoring and evaluation officers have to acquire the skills necessary for the conduct of complex evaluations. This is important in light of the thematic evaluations to be conducted by regional offices as contributions within the multi-year evaluation plan in the context of the MTSP, as well as in undertaking country programme evaluations.

E. Strengthening of New York headquarters

46. During 2001, the evaluation function at headquarters was re-engineered for the purpose of enabling UNICEF to use evaluation more strategically. In the context of the reorganization of the Programme Group, the Evaluation and Research Section of the EPP Division became the Evaluation Office, reporting to the Deputy Executive Director, Programme and Strategic Planning. The Office is now more independent and better positioned to contribute at the strategic level. The evaluation function at headquarters will focus on the country programme level and on the institutional management of the organization as a whole. For the latter purpose, the Evaluation Office has prepared a multi-year evaluation plan in the context of the MTSP. It is presented in paragraphs 50-57 below.

47. A senior-level Evaluation Committee will be created to deal with evaluation matters. It will be the formal forum that reviews evaluation reports and decides on the approval of the recommendations contained therein. The Evaluation Committee
will also review the annual follow-up reports on implementation of the recommendations. It will examine evaluation reports that have relevance at the global governance level. The reports produced by the Evaluation Office, as well as those produced by other headquarters divisions, will be reviewed. The Evaluation Committee will also review thematic evaluations conducted by the regional offices, as well as evaluations of country programmes of cooperation.

F. Fortifying management of the evaluation function

48. The evaluation function in UNICEF looks at activities undertaken or supported by UNICEF, examining their relevance, efficiency, effectiveness, impact and sustainability. Because of its important contribution to organizational learning, evaluation feedback is an integral part of the programme process. For the purpose of improving organizational learning and improving performance, in 2001, the Evaluation Office created a real-time, on-line Intranet access to the UNICEF organizational memory on performance, findings and lessons learned. The evaluation and research database is particularly tailored to the needs of UNICEF field offices. It allows users to access abstracts and full reports of evaluations and studies conducted by UNICEF and other organizations. It also serves as a reference source on methodological tools. In addition, the website allows electronic conferencing to foster professional exchange on performance assessment matters.

49. Another measure that will fortify the evaluation function is the approval of the competency profile for the different levels of evaluation positions, which will provide clearer technical criteria to select candidates. The competency profile will also be used to assess the training needs of present incumbents. The Evaluation Office will provide a technical assessment of the candidates. It will also maintain a network communication and exchange with evaluation officers, and provide them with updates on evaluation findings, events and methodologies on an ongoing basis.

V. Multi-year evaluation plan in support of the medium-term strategic plan

50. The MTSP seeks to combine a reinforced results-oriented management approach with a human rights-based approach to planning and programming. The MTSP establishes five organizational priorities; defines more clearly strategic objectives and indicators; and strengthens the strategic use of the evaluation function. The five organizational priorities are girls’ education; integrated early childhood development (ECD); immunization “plus”; fighting HIV/AIDS; and improved protection of children from violence, exploitation, abuse and discrimination. The strategies that UNICEF will use to pursue the organizational priorities include programme excellence; effective country programmes of cooperation; partnerships for shared success; influential information, communication and advocacy; and excellence in internal management and operations.

51. During the period of the MTSP, the evaluation function will focus on the country programme level and institutional management of the organization as a whole. It will look at the rationale, effectiveness, and administrative efficiency and economy of activities undertaken or supported by UNICEF. Thus, the organization
will enhance accountability and performance in terms of managing for results for the benefits of children. The organizational priorities of the MTSP will guide the parameters of the multi-year evaluation plan. Evaluations will be conducted with an emphasis on programmes, and organizational and policy considerations. Where possible and feasible, UNICEF will participate in joint evaluations with United Nations agencies and other partners. UNICEF will have opportunities to collaborate with the OECD/DAC evaluation group on thematic evaluations such as the current one on basic education. In the context of the Common Country Assessment/United Nations Development Assistance Framework (CCA/UNDAF), country programmes can be evaluated taking a United Nations system perspective. UNICEF can participate in multi-stakeholder evaluations such as those assessing the impact of sector-wide approaches. At a national level, UNICEF can contribute to thematic and sectoral evaluations involving the Government and other partners. On the basis of information needs for organizational decision-making, the types of contribution may range from desk reviews of existing evaluations and lessons learned to formal exercises involving stakeholders.

A. Evaluation of the organizational priorities

52. The five organizational priorities of the MTSP will guide the preparation of the annual global evaluation work plan. This annual global plan will incorporate the evaluation work led by headquarters, with contributions from the regions. At the end of the year, a summary of findings and lessons learned will be prepared and disseminated. Major findings will be incorporated in part II of the Executive Director’s annual report.

53. The following thematic evaluation activities are planned during the period of the MTSP:

2002-2003 HIV/AIDS
- Lessons learned from the evaluation of the Joint United Nations Programme on HIV/AIDS
- Methodology for assessing behavioural and institutional outcomes

Child protection
- Education as prevention against child labour

Immunization “plus”
- Evaluation of selected programmes

Integrated ECD
- Methodology for country case studies
- Baseline for the case studies

2004-2005 Integrated ECD
- Evaluation of ECD case studies
- Evaluation of Integrated Management of Childhood Illness case studies

HIV/AIDS
- Evaluation of behavioural and institutional outcomes

Girls’ education
- African Girls’ Education Initiative

Child protection
- Desk review of project review and lessons learned
B. Evaluation of the country programme of cooperation

54. Evaluation of the country programme of cooperation will become a systematized feature of the country programme process by the end of the four-year MTSP period. During the first two years of the MTSP, the Evaluation Office at headquarters, in cooperation with regional offices, will develop basic principles and methodologies, and will conduct a limited number of field tests. As of the third year of the MTSP, regional offices will assume full responsibility in this regard. The process will take into account the CCA/UNDAF and explore possibilities for the conduct of such exercises in this context. Tools for real-time evaluation of country programmes in the early crisis phase will also be developed and tested by the Evaluation Office, in collaboration with the Office of Internal Audit, the Office of Emergency Programmes and regional offices. The planned schedule of activities is as follows:

<table>
<thead>
<tr>
<th>Year</th>
<th>Evaluation of country programmes</th>
<th>Methodology and pilot cases</th>
<th>Methodology and testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002-2003</td>
<td>Evaluation of a country programme in a crisis situation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004-2005</td>
<td>Evaluation of a country programme in a crisis situation</td>
<td>Training and full introduction</td>
<td>Real-time evaluation to be used in the context of a major humanitarian crisis</td>
</tr>
</tbody>
</table>

C. Evaluation of organizational performance

55. The strategies used to implement the MTSP will guide the choice of functional and topical evaluations. Evaluation activities will be conducted for the purpose of assessing organizational performance in the context of excellence in internal management, advocacy and partnerships. In 2002, the Evaluation Office is conducting an evaluation of an information system (ChildInfo) and, in 2003, it will examine strategic considerations of the supply function.

D. Easier access to the organizational memory

56. During the MTSP period, the Evaluation Office, in collaboration with the UNICEF evaluators network, will improve the dissemination of monitoring and evaluation tools and findings from evaluation and research. In collaboration with the Division of Human Resources, an effort will be made to provide basic and advanced training in evaluation. In 2002, a web version of the training manual will be posted on the evaluation Intranet site. Over the 2002-2003 period, training sessions will be
offered in each region to ensure that each incumbent in an evaluation position meets the technical criteria, in accordance with the competency profile of the position.

57. The launching of the evaluation website last February enables UNICEF to provide access to the organizational memory of the evaluation and research database on the desktop or laptop of each UNICEF employee. UNICEF staff can now review and download evaluation tools and methodological references. Taking advantage of the reports contained in the evaluation and research database, desk reviews will be conducted to distil lessons learned by themes, sectors and topics related to the MTSP priorities. In addition, the UNICEF evaluation website provides links to all major evaluation websites. This is a priceless support tool made available to each country office.

VI. Draft recommendation

58. Evaluation activities conducted during the 1990s have had a noticeable impact on the quality of the organization’s work and thinking in those fields that were the major emphasis of past evaluation efforts. The challenge now is to ensure that evaluation efforts and results are given greater importance across all fields of activity and at all levels of management in a more systematic and strategic way.

59. Therefore, the Executive Director recommends that the Executive Board adopt the following draft recommendation:

The Executive Board

Endorses the “Report on the evaluation function in the context of the medium-term strategic plan” (E/ICEF/2002/10) as the official policy statement on the evaluation system of UNICEF.
United Nations Evaluation Group (UNEG)

Norms for Evaluation in the UN System

Towards a UN system better serving the peoples of the world; overcoming weaknesses and building on strengths from a strong evidence base

29 April 2005
Preamble

The United Nations system consists of various entities with diverse mandates and governing structures that aim to engender principles such as global governance, consensus building, peace and security, justice and international law, non-discrimination and gender equity, sustained socio-economic development, sustainable development, fair trade, humanitarian action and crime prevention. Above all, the UN system is collectively committed to furthering the Millennium Declaration.

The regulations that govern the evaluation of United Nations activities were promulgated on 19 April 2000 in the Secretary General’s bulletin¹. Similar regulations and policies have been issued in recent years in several UN system organizations. The United Nations Evaluation Group (UNEG), as a group of professional practitioners, undertook to define norms that aim at contributing to the professionalization of the evaluation function and at providing guidance to evaluation offices in preparing their evaluation policies or other aspects of their operations. This initiative was undertaken in part in response to General Assembly resolution A/RES/59/250² of December 2004, which encouraged UNEG to make further progress in a system-wide collaboration on evaluation, in particular the harmonization and simplification of methodologies, norms, standards and cycles of evaluation.

Resolutions of the General Assembly and governing bodies of UN organizations imply particular characteristics for the evaluation function within the United Nations system. Evaluation processes are to be inclusive, involving governments and other stakeholders. Evaluation activities require transparent approaches, reflecting intergovernmental collaboration. In addition, the General Assembly has requested that the UN system conducts evaluations in a way that fosters evaluation capacity building in member countries, to the extent that this is possible.

The norms seek to facilitate system-wide collaboration on evaluation by ensuring that evaluation entities within the UN follow agreed-upon basic principles. They provide a reference for strengthening, professionalizing and improving the quality of evaluation in all entities of the United Nations system, including funds, programmes and specialized agencies. The norms are consistent with other main sources³ and reflect the singularity of the United Nations system, characterized by its focus on people and respect for their rights, the importance of international values and principles, universality and neutrality,

³ These sources include, inter alia, the evaluation policies and guidelines existing within the various organizations of the United Nations system; OECD/DAC evaluation principles; national standards of OECD countries; evaluation policies of the international financial institutions; evaluation policies of the European Union; standards of evaluation associations; evaluation guidance developed by ALNAP for humanitarian action.
Norms for Evaluation in the UN System

its multiple stakeholders, its needs for global governance, its multidisciplinarity, and its complex accountability system. Last but not least, there is the challenge of international cooperation embedded in the Millennium Declaration and Development Goals.

To fulfil their mission of contributing to the greater effectiveness and the greater good of the world’s peoples, evaluation units within the UN system will strive for excellence and relevance by following the norms as outlined in this document.
Norms for Evaluation in the UN System

0 Introduction

0.1 The present document outlines the norms that are the guiding principles for evaluating the results achieved by the UN system, the performance of the organizations, the governing of the evaluation function within each entity of the UN system, and the value-added use of the evaluation function.

0.2 Complementary to these norms, a set of standards has been drawn from good practice of UNEG members. These will be revised from time to time and are intended to be applied as appropriate within each organization.

1 N1 - Definition

1.1 Purposes of evaluation include understanding why and the extent to which intended and unintended results are achieved, and their impact on stakeholders. Evaluation is an important source of evidence of the achievement of results and institutional performance. Evaluation is also an important contributor to building knowledge and to organizational learning. Evaluation is an important agent of change and plays a critical and credible role in supporting accountability.

1.2 An evaluation is an assessment, as systematic and impartial as possible, of an activity, project, programme, strategy, policy, topic, theme, sector, operational area, institutional performance, etc. It focuses on expected and achieved accomplishments, examining the results chain, processes, contextual factors and causality, in order to understand achievements or the lack thereof. It aims at determining the relevance, impact, effectiveness, efficiency and sustainability of the interventions and contributions of the organizations of the UN system. An evaluation should provide evidence-based information that is credible, reliable and useful, enabling the timely incorporation of findings, recommendations and lessons into the decision-making processes of the organizations of the UN system and its members.

1.3 Evaluation feeds into management and decision making processes, and makes an essential contribution to managing for results. Evaluation informs the planning, programming, budgeting, implementation and reporting cycle. It aims at improving the institutional relevance and the achievement of results, optimizing the use of

---

4 Hereinafter referred to as an “undertaking”.
5 This definition draws on Regulation 7.1 of Article VII of ST/SGB/2000/8 and from the widely accepted Principles for Evaluation of the Development Assistance Committee of the Organization for Economic Cooperation and Development (OECD DAC).
resources, providing client satisfaction and maximizing the impact of the contribution of the UN system.

1.4 There are other forms of assessment being conducted in the UN system. They vary in purpose and level of analysis, and may overlap to some extent. Evaluation is to be differentiated from the following:

a) **Appraisal**: a critical assessment of the potential value of an undertaking before a decision is made to implement it.

b) **Monitoring**: management’s continuous examination of progress achieved during the implementation of an undertaking to track compliance with the plan and to take necessary decisions to improve performance.

c) **Review**: the periodic or *ad hoc* often rapid assessments of the performance of an undertaking, that do not apply the due process of evaluation. Reviews tend to emphasize operational issues.

d) **Inspection**: a general examination that seeks to identify vulnerable areas and malfunctions and to propose corrective action.

e) **Investigation**: a specific examination of a claim of wrongdoing and provision of evidence for eventual prosecution or disciplinary measures.

f) **Audit**: an assessment of the adequacy of management controls to ensure the economical and efficient use of resources; the safeguarding of assets; the reliability of financial and other information; the compliance with regulations, rules and established policies; the effectiveness of risk management; and the adequacy of organizational structures, systems and processes.

g) **Research**: a systematic examination designed to develop or contribute to knowledge.

h) **Internal management consulting**: consulting services to help managers to implement changes that address organizational and managerial challenges and improve internal work processes.

1.5 Evaluation is not a decision-making process per se, but rather serves as an input to provide decision-makers with knowledge and evidence about performance and good practices. Although evaluation is used to assess undertakings, it should provide value-added for decision-oriented processes to assist in the improvement of present and future activities, projects, programmes, strategies and policies. Thus evaluation contributes to institutional policy-making, development effectiveness and organizational effectiveness.
1.6 There are many types of evaluations, such as those internally or externally-led, those adopting a summative or formative approach, those aimed at determining the attribution of an organization's own action or those performed jointly to assess collaborative efforts. An evaluation can be conducted in an ex-post fashion, at the end of phase, mid-point, at the terminal moment or real-time. The evaluation approach and method must be adapted to the nature of the undertaking to ensure due process and to facilitate stakeholder participation in order to support an informed decision-making process.

1.7 Evaluation is therefore about *Are we doing the right thing?* It examines the rationale, the justification of the undertaking, makes a reality check and looks at the satisfaction of intended beneficiaries. Evaluation is also about *Are we doing it right?* It assesses the effectiveness of achieving expected results. It examines the efficiency of the use of inputs to yield results. Finally, evaluation asks *Are there better ways of achieving the results?* Evaluation looks at alternative ways, good practices and lessons learned.

2 **N2 – Responsibility for Evaluation**

2.1 The Governing Bodies and/or the Heads of organizations in the UN system are responsible for fostering an enabling environment for evaluation and ensuring that the role and function of evaluation are clearly stated, reflecting the principles of the UNEG Norms for Evaluation, taking into account the specificities of each organization’s requirements.

2.2 The governance structures of evaluation vary. In some cases it rests with the Governing Bodies in others with the Head of the organization. Responsibility for evaluation should be specified in an evaluation policy.

2.3 The Governing Bodies and/or the Heads of organizations are also responsible for ensuring that adequate resources are allocated to enable the evaluation function to operate effectively and with due independence.

2.4 The Governing Bodies and/or Heads of organizations and of the evaluation functions are responsible for ensuring that evaluations are conducted in an impartial and independent fashion. They are also responsible for ensuring that evaluators have the freedom to conduct their work without repercussions for career development.

2.5 The Governing Bodies and/or Heads of organizations are responsible for appointing a professionally competent Head of the evaluation, who in turn is responsible for ensuring that the function is staffed by professionals competent in the conduct of evaluation.
2.6 The Governing Bodies and/or Heads of organizations and of the evaluation functions are responsible for ensuring that evaluation contributes to decision making and management. They should ensure that a system is in place for explicit planning for evaluation and for systematic consideration of the findings, conclusions and recommendations contained in evaluations. They should ensure appropriate follow-up measures including an action plan, or equivalent appropriate tools, with clear accountability for the implementation of the approved recommendations.

2.7 The Governing Bodies and/or Heads of organizations and of the evaluation functions are responsible for ensuring that there is a repository of evaluations and a mechanism for distilling and disseminating lessons to improve organizational learning and systemic improvement. They should also make evaluation findings available to stakeholders and other organizations of the UN system as well as to the public.

3 **N3 – Policy**

3.1 Each organization should develop an explicit policy statement on evaluation. The policy should provide a clear explanation of the concept, role and use of evaluation within the organization, including the institutional framework and definition of roles and responsibilities; an explanation of how the evaluation function and evaluations are planned, managed and budgeted; and a clear statement on disclosure and dissemination.

4 **N4 – Intentionality**

4.1 Proper application of the evaluation function implies that there is a clear intent to use evaluation findings. In the context of limited resources, the planning and selection of evaluation work has to be carefully done. Evaluations must be chosen and undertaken in a timely manner so that they can and do inform decision-making with relevant and timely information. Planning for evaluation must be an explicit part of planning and budgeting of the evaluation function and/or the organization as a whole. Annual or multi-year evaluation work programmes should be made public.

4.2 The evaluation plan can be the result of a cyclical or purposive selection of evaluation topics. The purpose, nature and scope of evaluation must be clear to evaluators and stakeholders. The plan for conducting each evaluation must ensure due process to ascertain the timely completion of the mandate, and consideration of the most cost-effective way to obtain and analyse the necessary information.
5 **N5 – Impartiality**

5.1 Impartiality is the absence of bias in due process, methodological rigour, consideration and presentation of achievements and challenges. It also implies that the views of all stakeholders are taken into account. In the event that interested parties have different views, these are to be reflected in the evaluation analysis and reporting.

5.2 Impartiality increases the credibility of evaluation and reduces the bias in the data gathering, analysis, findings, conclusions and recommendations. Impartiality provides legitimacy to evaluation and reduces the potential for conflict of interest.

5.3 The requirement for impartiality exists at all stages of the evaluation process, including the planning of evaluation, the formulation of mandate and scope, the selection of evaluation teams, the conduct of the evaluation and the formulation of findings and recommendations.

6 **N6 – Independence**

6.1 The evaluation function has to be located independently from the other management functions so that it is free from undue influence and that unbiased and transparent reporting is ensured. It needs to have full discretion in submitting directly its reports for consideration at the appropriate level of decision-making pertaining to the subject of evaluation.

6.2 The Head of evaluation must have the independence to supervise and report on evaluations as well as to track follow-up of management’s response resulting from evaluation.

6.3 To avoid conflict of interest and undue pressure, evaluators need to be independent, implying that members of an evaluation team must not have been directly responsible for the policy-setting, design, or overall management of the subject of evaluation, nor expect to be in the near future.

6.4 Evaluators must have no vested interest and have the full freedom to conduct impartially their evaluative work, without potential negative effects on their career development. They must be able to express their opinion in a free manner.

6.5 The independence of the evaluation function should not impinge the access that evaluators have to information on the subject of evaluation.

7 **N7 – Evaluability**
7.1 During the planning stage of an undertaking, evaluation functions can contribute to the process by improving the ability to evaluate the undertaking and by building an evaluation approach into the plan. To safeguard independence this should be performed in an advisory capacity only.

7.2 Before undertaking a major evaluation requiring a significant investment of resources, it may be useful to conduct an evaluability exercise. This would consist of verifying if there is clarity in the intent of the subject to be evaluated, sufficient measurable indicators, assessable reliable information sources and no major factor hindering an impartial evaluation process.

8 **N8 – Quality of Evaluation**

8.1 Each evaluation should employ design, planning and implementation processes that are inherently quality oriented, covering appropriate methodologies for data-collection, analysis and interpretation.

8.2 Evaluation reports must present in a complete and balanced way the evidence, findings, conclusions and recommendations. They must be brief and to the point and easy to understand. They must explain the methodology followed, highlight the methodological limitations of the evaluation, key concerns and evidenced-based findings, dissident views and consequent conclusions, recommendations and lessons. They must have an executive summary that encapsulates the essence of the information contained in the report, and facilitate dissemination and distillation of lessons.

9 **N9 - Competencies for Evaluation**

9.1 Each organization of the UN system should have formal job descriptions and selection criteria that state the basic professional requirements necessary for an evaluator and evaluation manager.

9.2 The Head of the evaluation function must have proven competencies in the management of an evaluation function and in the conduct of evaluation studies.

9.3 Evaluators must have the basic skill set for conducting evaluation studies and managing externally hired evaluators.

10 **N10 – Transparency and Consultation**

10.1 Transparency and consultation with the major stakeholders are essential features in all stages of the evaluation process. This improves the credibility and quality of the
Norms for Evaluation in the UN System

evaluation. It can facilitate consensus building and ownership of the findings, conclusions and recommendations.

10.2 Evaluation Terms of Reference and reports should be available to major stakeholders and be public documents. Documentation on evaluations in easily consultable and readable form should also contribute to both transparency and legitimacy.

11 N11 – Evaluation Ethics

11.1 Evaluators must have personal and professional integrity.

11.2 Evaluators must respect the right of institutions and individuals to provide information in confidence and ensure that sensitive data cannot be traced to its source. Evaluators must take care that those involved in evaluations have a chance to examine the statements attributed to them.

11.3 Evaluators must be sensitive to beliefs, manners and customs of the social and cultural environments in which they work.

11.4 In light of the United Nations Universal Declaration of Human Rights, evaluators must be sensitive to and address issues of discrimination and gender inequality.

11.5 Evaluations sometimes uncover evidence of wrongdoing. Such cases must be reported discreetly to the appropriate investigative body. Also, the evaluators are not expected to evaluate the personal performance of individuals and must balance an evaluation of management functions with due consideration for this principle.

12 N12 - Follow-up to Evaluation

12.1 Evaluation requires an explicit response by the governing authorities and management addressed by its recommendations. This may take the form of a management response, action plan and/or agreement clearly stating responsibilities and accountabilities.

12.2 There should be a systematic follow-up on the implementation of the evaluation recommendations that have been accepted by management and/or the Governing Bodies.

11.3 There should be a periodic report on the status of the implementation of the evaluation recommendations. This report should be presented to the Governing Bodies and/or the Head of the organization.
13 **N13 – Contribution to Knowledge Building**

13.1 Evaluation contributes to knowledge building and organizational improvement. Evaluations should be conducted and evaluation findings and recommendations presented in a manner that is easily understood by target audiences.

13.2 Evaluation findings and lessons drawn from evaluations should be accessible to target audiences in a user-friendly way. A repository of evaluation could be used to distil lessons that contribute to peer learning and the development of structured briefing material for the training of staff. This should be done in a way that facilitates the sharing of learning among stakeholders, including the organizations of the UN system, through a clear dissemination policy and contribution to knowledge networks.
United Nations Evaluation Group (UNEG)

Standards for Evaluation in the UN System

Towards a UN system better serving the peoples of the world; overcoming weaknesses and building on strengths from a strong evidence base

29 April 2005
Standards for Evaluation in the UN System

Preamble

The United Nations Evaluation Group (UNEG), as a group of professional practitioners, undertook to define norms and standards that aim at contributing to the professionalization of the evaluation function and at providing guidance to evaluation offices in preparing their evaluation policies or other aspects of their operations. This initiative was undertaken in part in response to General Assembly resolution A/RES/59/250 of December 2004, which encouraged UNEG to make further progress in a system-wide collaboration on evaluation, in particular the harmonization and simplification of methodologies, norms, standards and cycles of evaluation.

These standards build upon the Norms for Evaluation for the UN system. They are drawn from best practice of UNEG members. They are intended to guide the establishment of the institutional framework, management of the evaluation function, conduct and use of evaluations. They are also a reference for the competencies of evaluation practitioners and work ethics, and are intended to be applied as appropriate within each organization. UNEG will periodically update, elaborate and expand the coverage of these standards in the service of the UN system organizations.

2 In addition to evaluation policies and guidelines existing within the various organizations of the United Nations system, the standards have also drawn from the following sources: OECD/DAC evaluation principles; national standards of OECD countries; evaluation policies of the international financial institutions; evaluation policies of the European Union; standards of evaluation associations; evaluation guidance developed by ALNAP for humanitarian action.
3 UN organizations refer hereinafter to all organizations, funds and programmes as well as specialized agencies of the UN system.
Institutional Framework

**Standard 1.1:** United Nations organizations should have an adequate institutional framework for the effective management of their evaluation function.

1. A comprehensive institutional framework for the management of the evaluation function and conduct of evaluations is crucial to ensure an effective evaluation process.

2. Such an institutional framework should address the following requirements:
   - Provide institutional and high-level management understanding of and support for the evaluation function's key role in contributing to the effectiveness of the organization.
   - Ensure that evaluation is part of the organization’s governance and management functions. Evaluation makes an essential contribution to managing for results.
   - Promote a culture that values evaluation as a basis for learning.
   - Facilitate an independent and impartial evaluation process by ensuring that the evaluation function is independent of other management functions. The Head of evaluation should report directly to the Governing Body of the organization or the Head of the organization.
   - Ensure adequate financial and human resources for evaluation in order to allow efficient and effective delivery of services by a competent evaluation function and enable evaluation capacity strengthening.
   - Encourage partnerships and cooperation on evaluation within the UN system, as well as with other relevant institutions.

**Standard 1.2:** UN organizations should develop an evaluation policy and regularly update it, taking into account the Norms and Standards for Evaluation in the UN system.

3. The evaluation policy should be approved by the Governing Bodies of the organizations and/or Head of the organization, and should be in line with the applicable UNEG Norms for Evaluation, and with organizational corporate goals and strategies. The evaluation policy should include:
Standards for Evaluation in the UN System

- clear explanation of the concept and role of evaluation within the organization;
- clear definition of the roles and responsibilities of the evaluation professionals, senior management and programme managers;
- an emphasis on the need for adherence to the organization's evaluation guidelines;
- explanation of how evaluations are prioritized and planned;
- description of how evaluations are organized, managed and budgeted;
- an emphasis on the requirements for the follow-up of evaluations;
- clear statement on disclosure and dissemination.

Standard 1.3: UN organizations should ensure that evaluation plans of evaluation activities are submitted to their Governing Bodies and/or Heads of organizations for review and/or approval.

4. The Governing Bodies and/or the Head of the organization should receive not only the evaluation plan, but also a progress report on the implementation of both the evaluation plan as well as the recommendations emanating from the evaluations.

Standard 1.4: UN organizations should ensure appropriate evaluation follow-up mechanisms and have an explicit disclosure policy.

5. Appropriate evaluation follow-up mechanisms should exist within the organization, ensuring that evaluation recommendations are properly utilized and implemented in a timely fashion and that evaluation findings are linked to future activities.

6. A disclosure policy should ensure the transparent dissemination of evaluation results, including making reports broadly available to the Governing Bodies and the public, except in those cases where the reasonable protection and confidentiality of some stakeholders is required.
Management of the Evaluation Function

Standard 1.5: The Head of evaluation has a lead role in ensuring that the evaluation function is fully operational and that evaluation work is conducted according to the highest professional standards.

7. Within the comprehensive institutional framework, the management of the evaluation function, entrusted to the Head of evaluation, should ensure that:

- an evaluation policy is developed and regularly updated;
- the budget for evaluations is managed in an efficient manner;
- an evaluation plan of evaluation activities is developed as part of the organization's planning and budgeting cycle, on an annual or biannual basis. The plan should prioritize those areas most in need of evaluation, and specify adequate resources for the planning, conduct and follow-up of evaluations;
- adequate evaluation methodologies are adopted, developed and updated frequently;
- the evaluations are conducted according to defined quality standards, in a timely manner, in order to serve as a useful tool for the intended stakeholders/users;
- reporting to high-level management is timely and relevant to their needs, thereby supporting an informed management and policy decision-making process;
- regular progress reports are compiled on the implementation of the evaluation plan and/or the implementation of the recommendations emanating from the evaluations already carried out, to be submitted to the Governing Bodies and/or Heads of organizations;
- lessons from evaluations are distilled and disseminated as appropriate.

Standard 1.6: The Head of evaluation is responsible for ensuring the preparation of evaluation guidelines.

8. Evaluation guidelines should be prepared and include the following:

- evaluation methodologies that should reflect the highest professional standards;
- evaluation processes, ensuring that evaluations are conducted in an objective, impartial, open and participatory manner, based on empirically verified evidence that is valid and reliable, with results being made available;
Standards for Evaluation in the UN System

- ethics, ensuring that evaluations are carried out with due respect and regard to those being evaluated.

Standard 1.7: The Head of evaluation should ensure that the evaluation function is dynamic, adapting to new developments and changing needs both within and outside the organization.

9. In particular the management of the evaluation function should include:
   - raising awareness and/or building evaluation capacity;
   - facilitation and management of evaluation networks;
   - design and implementation of evaluation methodologies and systems;
   - ensuring the maintenance of institutional memory of evaluations through user-friendly mechanisms;
   - promoting the compilation of lessons in a systematic manner.

2. Competencies and Ethics

1. All those engaged in designing, conducting and managing evaluation activities should aspire to conduct high quality and ethical work guided by professional standards and ethical and moral principles.

Competencies

Standard 2.1: Persons engaged in designing, conducting and managing evaluation activities should possess core evaluation competencies.

2. Evaluation competencies refer to the qualifications, skills, experience and attributes required by those employed within the evaluation function to carry out their duties as stipulated and to ensure the credibility of the process.

3. Competencies are required for all those engaged in designing, conducting and managing evaluation activities, managing evaluators, conducting training and capacity development and designing and implementing evaluation methodologies and systems.

4. Some skills are particularly useful for persons conducting evaluations as “evaluators”, while others are needed for persons who manage evaluations as “evaluation managers”. The term “evaluators” used below encompasses both roles.
5. Evaluators should declare any conflict of interest to clients before embarking on an evaluation project, and at any point where such conflict occurs. This includes conflict of interest on the part of either the evaluator or the stakeholder.

6. Evaluators should accurately represent their level of skills and knowledge. Similarly, evaluators should practice within the limits of their professional training and competence, and should decline to conduct evaluations that fall substantially outside those limits.

<table>
<thead>
<tr>
<th>Standard 2.2: Evaluators should have relevant educational background, qualification and training in evaluation.</th>
</tr>
</thead>
</table>

7. Evaluators should preferably have an advanced university degree or equivalent background in social sciences or other relevant disciplines, with specialized training in areas such as evaluation, project management, social statistics, advanced statistical research and analysis.

8. Evaluators should continually seek to maintain and improve their competencies in order to provide the highest level of performance in their evaluations. This continuing professional development might include formal seminars and workshops, self-study, evaluations of one's own practice, and working with other evaluators to learn from their skills and expertise.

<table>
<thead>
<tr>
<th>Standard 2.3: Evaluators should have professional work experience relevant to evaluation.</th>
</tr>
</thead>
</table>

9. Evaluators should also have relevant professional experience in:
   - design and management of evaluation processes, including with multiple stakeholders;
   - survey design and implementation;
   - social science research;
   - project/programme/policy planning, monitoring and management.

<table>
<thead>
<tr>
<th>Standard 2.4: Evaluators need to have specific technical knowledge of, and be familiar with, the methodology or approach that will be needed for the specific evaluation to be undertaken, as well as certain managerial and personal skills.</th>
</tr>
</thead>
</table>

10. Specialized experience and/or methodological/technical knowledge, including some specific data collection and analytical skills, may be particularly useful in the following areas:
   - understanding of human rights-based approaches to programming;
- understanding of gender considerations;
- understanding of Results Based Management (RBM) principles;
- logic modelling/logical framework analysis;
- real-time, utilization-focused, joint, summative and formative evaluation;
- quantitative and qualitative data collection and analysis;
- rapid assessment procedures;
- participatory approaches.

11. The evaluator, whose responsibilities include the management of evaluation, needs specific managerial skills:
- management of evaluation process;
- planning, setting standards and monitoring work;
- management of human and financial resources;
- team leadership;
- strategic and global thinking;
- foresight and problem solving.

12. The evaluator also needs certain personal skills that are particularly useful in evaluation:
- team work and cooperation;
- capability to bring together diverse stakeholders;
- communication;
- strong drafting skills;
- analytical skills;
- negotiation skills;
- language skills adapted to the region where the evaluation takes place.

**Ethics**

| Standard 2.5: Evaluators should be sensitive to beliefs, manners and customs and act with integrity and honesty in their relationships with all stakeholders. |

13. In line with the UN Universal Declaration of Human Rights and other human rights conventions, evaluators should operate in accordance with international values.

14. Evaluators should be aware of differences in culture, local customs, religious beliefs and practices, personal interaction and gender roles, disability, age and ethnicity, and be mindful of the potential implications of these differences when planning, carrying out and reporting on evaluations.

15. Evaluators must ensure the honesty and integrity of the entire evaluation process. Evaluators also have an overriding responsibility to ensure that evaluation activities are independent, impartial and accurate.
Standard 2.6: Evaluators should ensure that their contacts with individuals are characterized by respect.

16. Evaluators should avoid offending the dignity and self-respect of those persons with whom they come in contact in the course of the evaluation.

17. Knowing that evaluation might often negatively affect the interests of some stakeholders, evaluators should conduct the evaluation and communicate its purpose and results in a way that clearly respects the stakeholders' dignity and self-worth.

Standard 2.7: Evaluators should protect the anonymity and confidentiality of individual informants.

18. Evaluators should provide maximum notice, minimize demands on time, and respect people’s right to privacy.

19. Evaluators must respect people’s right to provide information in confidence, and must ensure that sensitive information cannot be traced to its source. They should also inform participants about the scope and limits of confidentiality.

20. Evaluators are not expected to evaluate individuals, and must balance an evaluation of management functions with this general principle.

21. Evaluators have a responsibility to note issues and findings that may not relate directly to the Terms of Reference. They should consult with other relevant oversight entities when there is any doubt about if and how issues, such as evidence of wrongdoing, should be reported.

Standard 2.8: Evaluators are responsible for their performance and their product(s).

22. Evaluators are responsible for the clear, accurate and fair written and/or oral presentation of study limitations, findings and recommendations.

23. Evaluators should be responsible for the completion of the evaluation within a reasonably planned time, acknowledging unprecedented delays resulting from factors beyond the evaluator's control.
Standards for Evaluation in the UN System

3. Conducting Evaluations

Design

Standard 3.1: The evaluation should be designed to ensure timely, valid and reliable information that will be relevant for the subject being assessed.

1. The conduct of evaluations follows the cyclical planning at various levels, which is comprised of different stages: planning, design, implementation and follow-up.

Standard 3.2: The Terms of Reference should provide the purpose and describe the process and the product of the evaluation.

2. The design of an evaluation should be described as precisely as possible in the Terms of Reference, which should include the following elements:
   - context for the evaluation;
   - purpose of the evaluation;
   - scope (outlining what is covered and what is not covered by the evaluation);
   - evaluation criteria (inter alia relevance, efficiency, effectiveness, impact, sustainability);
   - key evaluation questions;
   - methodology – approach for data collection and analysis and involvement of stakeholders;
   - workplan, organization and budget;
   - products and reporting;
   - use of evaluation results, including responsibilities for such use.

Standard 3.3: The purpose and context of the evaluation should be clearly stated, providing a specific justification for undertaking the evaluation at a particular point in time.

3. The purpose of the evaluation must be clearly and accurately defined bearing in mind the main information needs of the intended users of the evaluation. The purpose discusses why the evaluation is being done, what triggered it and how it will be used. The purpose also relates to the timing of the evaluation at various junctions in the management cycle. This adds to the clarity of the evaluation and should provide the broad orientation, which is then further elaborated in the objectives and scope of the evaluation.
Standards for Evaluation in the UN System

Standard 3.4: The subject to be evaluated should be clearly described.

4. The subject to be evaluated should be described in terms of what it aims to achieve, how the designers thought that it would address the problem they had identified, implementation modalities, and any intentional, or unintentional, change in implementation.

5. Other elements include the importance or parameters of the subject to be evaluated including its cost and its relative weight with respect, for example, to the organization’s overall activities. At the very least, the description should include the number of participants/people reached by the undertaking.

Standard 3.5: Evaluation objectives should be realistic and achievable, in light of the information that can be collected in the context of the undertaking. The scope of the evaluation also needs to be clearly defined.

6. The objectives of the evaluation should follow from the purpose of the evaluation. They should be clear and agreed upon by all stakeholders involved.

7. Scope determines the boundaries of the evaluation, tailoring the objectives and evaluation criteria to the given situation. It should also make the coverage of the evaluation explicit (time period, phase in implementation, geographical area and the dimensions of stakeholder involvement being examined). The limits of the evaluation should also be acknowledged within the scope.

8. Evaluations may also be oriented by evaluation questions. These add more detail to the objectives and contribute to defining the scope.

9. The objectives and scope of the evaluation are critical references to determine the evaluation methodology and required resources.

Standard 3.6: The evaluation design should clearly spell out the evaluation criteria against which the subject to be evaluated will be assessed.

10. The most commonly applied evaluation criteria are the following: relevance, efficiency, effectiveness, impact, value-for-money, client satisfaction and sustainability. Criteria for humanitarian response should also include: coverage, coordination, coherence, connectedness and protection. Not all criteria are applicable to every evaluation.
Standards for Evaluation in the UN System

Standard 3.7: Evaluation methodologies should be sufficiently rigorous to assess the subject of evaluation and ensure a complete, fair and unbiased assessment.

11. The evaluation methodologies to be used for data collection, analysis and involvement of stakeholders should be appropriate to the subject to be evaluated, to ensure that the information collected is valid, reliable and sufficient to meet the evaluation objectives, and that the assessment is complete, fair and unbiased.

12. Evaluation methods depend on the information sought, and the type of data being analysed. The data should come from a variety of sources to ensure its accuracy, validity and reliability, and that all affected people/stakeholders are considered. Methodology should explicitly address issues of gender and under-represented groups.

13. The limitations of the chosen evaluation methods should also be acknowledged.

Standard 3.8: An evaluation should assess cost effectiveness, to the extent feasible.

14. Using a range of cost analysis approaches, from the elaborate cost-effectiveness and cost-benefit analysis, to cost-efficiency analysis, to a quick cost comparison, an evaluation should, to the extent possible, pursue the following broad questions:
   - How do actual costs compare to other similar benchmarks?
   - What is the cheapest or most efficient way to get the expected results?
   - What are the cost implications of scaling up or down?
   - What are the costs of replicating the subject being evaluated in a different environment?
   - Is the subject being evaluated worth doing? Do economic benefits outweigh the costs?
   - How do costs affect the sustainability of the results?

15. Cost analysis in evaluation builds on financial information, but may also involve calculating “economic costs” such as human resources, labour-in-kind, opportunity costs, etc.

16. The scope of cost analysis, i.e. whether cost comparison is made concerning impacts, outcomes or outputs, will depend on the purpose of the evaluation and the evaluation questions posed. Cost analysis must be explicit in terms of the different perspectives from which costs are analysed (donors, a single organization, primary stakeholders) and the limitations – the complexity of the subject (multiple programme objectives, partners, financial systems), the availability of data and the time and resources invested.

17. Cost analysis is not always feasible. Where no cost analysis is included in an evaluation, some rationale for this exclusion should be included in the objectives or methodology section.
18. It is expected that evaluators point out areas of obvious inefficient use of resources.

Standard 3.9: The evaluation design should, when relevant, include considerations as to what extent the UN system’s commitment to the human rights-based approach has been incorporated in the design of the undertaking to be evaluated.

19. UN organizations are guided by the United Nations Charter, and have a responsibility and mission to assist Member States to meet their obligations towards the realization of the human rights of those who live within their jurisdiction. Human rights treaties, mechanisms and instruments provide UN entities with a guiding frame of reference and a legal foundation for ethical and moral principles, and should guide evaluation work. Consideration should also be given to gender issues and hard-to-reach and vulnerable groups.

20. The evaluation design might in addition include some process of ethical review of the initial design of the undertaking to be evaluated, including:

- the balance of cost and benefits to participants including potential negative impact;
- the ethics of who is included and excluded in the evaluation and how this is done;
- handling of privacy and confidentiality;
- practices of obtaining informed consent;
- feedback to participants;
- mechanisms for shaping and monitoring the behaviour and practice of evaluators and data collectors.

Process

Standard 3.10: The relationship between the evaluator and the commissioner(s) of an evaluation must, from the outset, be characterized by mutual respect and trust.

21. The responsibilities of the parties who agree to conduct an evaluation (specifying what, how, by whom, and when what is to be done) should be set forth in a written agreement in order to obligate the contracting parties to fulfil all the agreed upon conditions, or if not, to renegotiate the agreement. Agreements, such as Terms of Reference, should be established at least in the following areas: financing, time frame, persons involved, reports to be produced or published, content, methodology, and procedures to be followed. Such an agreement reduces the likelihood that misunderstandings will arise between the contracting parties and makes it easier to
resolve them if they do arise. Providing an inception report at the start of the evaluation is a useful way of formalizing such an agreement and ensuring proper interpretation of the Terms of Reference.

22. Evaluators should consult with the commissioner(s) of the evaluation on contractual decisions such as confidentiality, privacy, communication, and ownership of findings and reports.

**Standard 3.11:** Stakeholders should be consulted in the planning, design, conduct and follow-up of evaluations.

23. Stakeholders must be identified and consulted when planning the evaluation (key issues, method, timing, responsibilities) and should be kept informed throughout the evaluation process. The evaluation approach must consider learning and participation opportunities (e.g. workshops, learning groups, debriefing, participation in the field visits) to ensure that key stakeholders are fully integrated into the evaluation learning process.

24. When feasible, a core learning group or steering group composed of representatives of the various stakeholders in the evaluation may be created. This group’s role is to act as a sounding board, facilitate and review the work of the evaluation. In addition, this group may be tasked with facilitating the dissemination and application of the results and other follow-up action.

**Standard 3.12:** A peer review, or reference group, composed of external experts may be particularly useful.

25. Depending on the scope and complexity of the evaluation, it may be useful to establish a peer review or reference group composed of experts in the technical topics covered by the evaluation. This group would provide substantive guidance to the evaluation process (e.g. provide inputs on the Terms of Reference and provide quality control of the draft report).

**Selection of Team**

**Standard 3.13:** Evaluations should be conducted by well-qualified evaluation teams.

26. The number of evaluators in a given team depends on the size of the evaluation. Multi-faceted evaluations need to be undertaken by multi-disciplinary teams.

27. Evaluators should be selected on the basis of competence, and by means of a transparent process.
28. The members selected must bring different types of expertise and experience to the team. If possible, at least one member of the team should be experienced in the sector or technical areas addressed by the evaluation, or have a sound knowledge of the subject to be evaluated. At least one other should preferably be an evaluation specialist and be experienced in using the specific evaluation methodologies that will be employed for that evaluation. The evaluation team should also possess a broad knowledge and understanding of the major economic and social development issues and problems in the country(ies) where the evaluation is taking place or in similar countries in the region. Background or familiarity with emergency situations may also be required, both for the conduct of the exercise itself, and for understanding the particular context of the evaluation.

Standard 3.14: The composition of evaluation teams should be gender balanced, geographically diverse and include professionals from the countries or regions concerned.

29. Qualified, competent and experienced professional firms or individuals from concerned countries should be involved, whenever possible, in the conduct of evaluations, in order, inter alia, to ensure that national/local knowledge and information is adequately taken into account in evaluations and to support evaluation capacity building in developing countries. The conduct of evaluations may also be out-sourced to national private sector and civil society organizations. Joint evaluations with governments or other stakeholders should equally be encouraged.

30. Members of the evaluation team should also familiarize themselves with the cultural and social values and characteristics of the recipients and intended beneficiaries. In this way, they will be better equipped to understand and respect local customs, beliefs and practices throughout the evaluation work.

**Implementation**

Standard 3.15: Evaluations should be conducted in a professional and ethical manner.

31. Evaluations should be carried out in a participatory and ethical manner and the welfare of the stakeholders should be given due respect and consideration (human rights, dignity and fairness). Evaluations must be gender and culturally sensitive and respect the confidentiality, protection of source and dignity of those interviewed.

32. Evaluation procedures should be conducted in a realistic, diplomatic, cost-conscious and cost-effective manner.
33. Evaluations must be accurate and well-documented and deploy transparent methods that provide valid and reliable information. Evaluation team members should have an opportunity to disassociate themselves from particular judgments and recommendations. Any unresolved differences of opinion within the team should be acknowledged in the report.

34. Evaluations should be conducted in a complete and balanced manner so that the different perspectives are addressed and analysed. Key findings must be substantiated through triangulation. Any conflict of interest should be addressed openly and honestly so that it does not undermine the evaluation outcome.

35. Evaluators should discuss, in a contextually appropriate way, those values, assumptions, theories, methods, results, and analyses that significantly affect the interpretation of the evaluative findings. These statements apply to all aspects of the evaluation, from its initial conceptualization to the eventual use of findings.

36. The rights and well-being of individuals should not be affected negatively in planning and carrying out an evaluation. This needs to be communicated to all persons involved in an evaluation, and its foreseeable consequences for the evaluation discussed.

**Reporting**

```
Standard 3.16: The final evaluation report should be logically structured, containing evidence-based findings, conclusions, lessons and recommendations, and should be free of information that is not relevant to the overall analysis. The report should be presented in a way that makes the information accessible and comprehensible.
```

37. A reader of an evaluation report must be able to understand:

- the purpose of the evaluation;
- exactly what was evaluated;
- how the evaluation was designed and conducted;
- what evidence was found;
- what conclusions were drawn;
- what recommendations were made;
- what lessons were distilled.

38. If evaluators identify fraud, misconduct, abuse of power and rights violation, they should confidentially refer the matter to the appropriate UN authorities to investigate such matters. Evaluations should not substitute, or be used for, decision-making in individual human resources matters.

39. Evaluators should allow all relevant stakeholders to have access to appropriate evaluative information, and should actively disseminate that information to stakeholders.
Standards for Evaluation in the UN System

if possible. Communications to a given stakeholder should always include all important results that may bear on the interests of that stakeholder. In all cases, evaluators should strive to present results as clearly and simply as possible so that clients and other stakeholders can easily understand the evaluation process and results.

Follow-up

Standard 3.17: Evaluation requires an explicit response by the governing authorities and management addressed by its recommendations.

40. As per the Norms, this may take the form of a management response, action plan and/or agreement clearly stating responsibilities and accountabilities.

41. Follow-up on the implementation of the evaluation recommendations that have been accepted by management and/or the Governing Bodies should be systematically carried out.

42. Periodic reporting on the status of the implementation of the evaluation recommendations should also be conducted. This report should be presented to the Governing Bodies and/or the Head of the organization.

4. Evaluation Reports

Standard 4.1: The title page and opening pages should provide key basic information.

1. The following information should be easily accessible in the first few pages of the report:
   - name of the subject (i.e. activity, programme, policy etc.) being evaluated;
   - date;
   - table of contents, including annexes;
   - name and organization(s) of the evaluators;
   - name and address of the organization(s) that commissioned the evaluation.

Standard 4.2: The evaluation report should contain an Executive Summary.

2. An Executive Summary should provide a synopsis of the substantive elements of the evaluation report. To facilitate higher readership, the Executive Summary should be short, two to three pages, and should “stand alone”. The level of information should
provide the uninitiated reader with a clear understanding of what was found and recommended and what has been learned from the evaluation.

3. The Executive Summary should include:
   - a brief description of the subject being evaluated;
   - the context, present situation, and description of the subject vis-à-vis other related matters;
   - the purpose of the evaluation;
   - the objectives of the evaluation;
   - the intended audience of the report;
   - a short description of methodology, including rationale for choice of methodology, data sources used, data collection and analysis methods used, and major limitations;
   - the most important findings and conclusions;
   - main recommendations.

Standard 4.3: The subject being evaluated should be clearly described, including the logic model and/or the expected results chain and intended impact, its implementation strategy and key assumptions.

4. The evaluation report should clearly describe what the purpose of the subject being evaluated is and how the designers thought it would address the identified problem. Additional important elements include: the importance, scope and scale of the subject being evaluated; a description of the recipients / intended beneficiaries and stakeholders; and budget figures.

5. The description of the subject being evaluated should be as short as possible while ensuring that all pertinent information is provided. If additional details are deemed necessary, a description including the logic model can be provided in an annex.

Standard 4.4: The role and contributions of the UN organizations and other stakeholders to the subject being evaluated should be clearly described.

6. The report should describe who is involved, their roles and their contributions to the subject being evaluated, including financial resources, in-kind contributions, technical assistance, participation, staff time, training, leadership, advocacy, lobbying, and any contributions from primary stakeholders, such as communities. An attempt should be made to clarify what partners contributed to which outcome.

7. Users will want to compare this with who was involved in the evaluation to assess how different points of view were included.
Standards for Evaluation in the UN System

Standard 4.5: The purpose and context of the evaluation should be described.

8. The purpose should discuss why the evaluation is being done, how it will be used and what decisions will be taken after the evaluation is complete. The context should be described in order to provide an understanding of the setting in which the evaluation took place.

Standard 4.6: The evaluation report should provide an explanation of the evaluation criteria that were used by the evaluators.

9. Not all criteria are applicable to every evaluation. The rationale for not using a particular criterion should be explained in the report, as should any limitations in applying the evaluation criteria. Performance standards or benchmarks used in the evaluation should also be described.

10. It is important to make the basis of value judgments transparent.

Standard 4.7: The evaluation report should provide a clear explanation of the evaluation objectives as well as the scope of the evaluation.

11. The original objectives of the evaluation should be described, as well as any changes made to the evaluation design.

12. The scope of the evaluation should be described, making the coverage of the evaluation explicit. The limits of the evaluation should also be acknowledged.

13. The original evaluation questions should be explained, as well as those that were added during the evaluation. These are critical references against which the content of the report ought to be compared to.

14. The objectives and scope of the evaluation are also critical references to judge whether the methodology selected and resources allocated were adequate.

Standard 4.8: The evaluation report should indicate the extent to which gender issues and relevant human rights considerations were incorporated where applicable.

15. The evaluation report should include a description of, *inter alia*:

- how gender issues were implemented as a cross-cutting theme in programming, and if the subject being evaluated gave sufficient attention to promote gender equality and gender-sensitivity;
Standards for Evaluation in the UN System

- whether the subject being evaluated paid attention to effects on marginalized, vulnerable and hard-to-reach groups;
- whether the subject being evaluated was informed by human rights treaties and instruments;
- to what extent the subject being evaluated identified the relevant human rights claims and obligations;
- how gaps were identified in the capacity of rights-holders to claim their rights, and of duty-bearers to fulfil their obligations, including an analysis of gender and marginalized and vulnerable groups, and how the design and implementation of the subject being evaluated addressed these gaps;
- how the subject being evaluated monitored and viewed results within this rights framework.

Standard 4.9: The applied evaluation methodology should be described in a transparent way, including any limitations to the methodology.

16. A comprehensive, but not excessive, description of the critical aspects of methodology should be contained in the evaluation report to allow the user(s) of the evaluation to come to their own conclusions about the quality of the data. Any description of the methodology should include:

- data sources;
- description of data collection methods and analysis (including level of precision required for quantitative methods, value scales or coding used for qualitative analysis);
- description of sampling (area and population to be represented, rationale for selection, mechanics of selection, numbers selected out of potential subjects, limitations to sample);
- reference indicators and benchmarks, where relevant (previous indicators, national statistics, etc.);
- evaluation team, including the involvement of individual team members;
- the evaluation plan;
- key limitations.

The annexes should include the following:

- more detail on any of the above;
- data collection instruments (surveys, checklists, etc.);
- system for ensuring data quality through monitoring of data collection and oversight;
- a more detailed discussion of limitations as needed.

Standard 4.10: The evaluation should give a complete description of stakeholders’ participation.
17. The level of participation of stakeholders in the evaluation should be described, including the rationale for selecting that particular level. While not all evaluations can be participatory to the same degree, it is important that consideration is given to participation of stakeholders, as such participation is increasingly recognized as a critical factor in the use of conclusions, recommendations and lessons. A human rights-based approach to programming adds emphasis to the participation of primary stakeholders. In many cases, this clearly points to the involvement of people and communities. Also, including certain groups of stakeholders may be necessary for a complete and fair assessment.

Standard 4.11: The evaluation report should include a discussion of the extent to which the evaluation design included ethical safeguards where appropriate.

18. The report should have a good description of ethical considerations, including the rationale behind the evaluation design and the mechanisms to protect participants where appropriate. This includes protection of the confidentiality, dignity, rights and welfare of human subjects, including children, and respect for the values of the beneficiary communities.

Standard 4.12: In presenting the findings, inputs, outputs, and outcomes / impacts should be measured to the extent possible (or an appropriate rationale given as to why not).

19. Findings regarding inputs for the completion of activities or process achievements should be distinguished clearly from outputs, outcomes and impact.

20. Outcomes and impacts should include any unintended effects, whether beneficial or harmful. Additionally, any multiplier or downstream effects of the subject being evaluated should be included. To the extent possible, each of these should be measured either quantitatively or qualitatively. In using such measurements, benchmarks should be referred to.

21. The report should make a logical distinction in the findings, showing the progression from implementation to results with an appropriate measurement and analysis of the results chain, or a rationale as to why an analysis of results was not provided.

22. Data does not need to be presented in full; only data that supports a finding needs to be given, and full data can be put in an annex. Additionally, reports should not segregate findings by data source.

23. Findings should cover all of the evaluation objectives and use the data collected.
Standards for Evaluation in the UN System

Standard 4.13: Analysis should include appropriate discussion of the relative contributions of stakeholders to results.

24. Results attributed to the subject being evaluated should be related back to the contributions of different stakeholders. There should be a sense of proportionality between the relative contributions of each, and the results observed. This is an integral element of accountability to partners, donors and primary stakeholders.

25. If such an analysis is not included in the report, the reason why it was not done should be clearly indicated. For instance, if an evaluation is done early in the management cycle, results or any link to a stakeholder’s contribution may not be found.

Standard 4.14: Reasons for accomplishments and difficulties of the subject being evaluated, especially constraining and enabling factors, should be identified to the extent possible.

26. An evaluation report should go beyond a mere description of implementation and outcomes and include an analysis, based on the findings, of the underlying causes, constraints, strengths on which to build on, and opportunities. External factors contributing to the accomplishments and difficulties should be identified and analysed to the extent possible, including the social, political or environmental situation.

27. An explanation of context contributes to the utility and accuracy of the evaluation. An understanding of which external factors contributed to the success or failure of a subject being evaluated helps determine how such factors will affect the future of the subject being evaluated, or whether it could be replicated elsewhere.

Standard 4.15: Conclusions need to be substantiated by findings consistent with data collected and methodology, and represent insights into identification and/or solutions of important problems or issues.

28. Conclusions should add value to the findings. The logic behind conclusions and the correlation to actual findings should be clear.

29. Conclusions must focus on issues of significance to the subject being evaluated, determined by the evaluation objectives and the key evaluation questions. Simple conclusions that are already well known and obvious are not useful, and should be avoided.

30. Conclusions regarding attribution of results, which are most often tentative, require clear detailing of what is known and what can plausibly be assumed in order to make the logic from findings to conclusions more transparent, and thereby increase the credibility of the conclusions.
Standards for Evaluation in the UN System

Standard 4.16: Recommendations should be firmly based on evidence and analysis, be relevant and realistic, with priorities for action made clear.

31. For accuracy and credibility, recommendations should be the logical implications of the findings and conclusions. Recommendations should also be relevant to the subject being evaluated, the Terms of Reference and the objectives of the evaluation, and should be formulated in a clear and concise manner. Additionally, recommendations should be prioritized to the extent possible.

32. Recommendations should state responsibilities and the time frame for their implementation, to the extent possible.

Standard 4.17: Lessons, when presented, should be generalized beyond the immediate subject being evaluated to indicate what wider relevance they might have.

33. Not all evaluations generate lessons. Lessons should only be drawn if they represent contributions to general knowledge. They should be well supported by the findings and conclusions of the evaluation. They may refine or add to commonly accepted lessons, but should not be merely a repetition of common knowledge.

34. A good evaluation report has correctly identified lessons that stem logically from the findings, presents an analysis of how they can be applied to different contexts and/or different sectors, and takes into account evidential limitations such as generalizing from single point observations.

Standard 4.18: Annexes should be complete and relevant.

35. Additional supplementary information to the evaluation that should be included in annexes includes:
   - list of persons interviewed (if confidentiality allows) and sites visited;
   - data collection instruments (copies of questionnaires, surveys, etc.);
   - the original Terms of Reference for the evaluation;
   - list of abbreviations.

36. The annexes increase the usability and the credibility of the report.
Evaluation standards

The following are the programme evaluation standards developed by the American Joint Committee on Standards for Educational Evaluation (AJCSEE) which have increasingly been promoted through professional evaluation associations, including the American and African evaluation associations. The African Evaluation Association has further adapted the original AJCSEE standards. Regardless of the version adopted, these standards can be used both as a guide for managing the evaluation process and to assess an existing evaluation. The standards highlight the considerations that must be weighed in formulating an evaluation design.

- **Utility**: seek to ensure that an evaluation will serve the information needs of intended users.
- **Feasibility**: seek to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.
- **Propriety**: seek to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.
- **Accuracy**: seek to ensure that an evaluation will reveal and convey technically adequate information about the features that determine the worth or merit of the programme being evaluated.

<table>
<thead>
<tr>
<th>Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Stakeholder Identification</td>
</tr>
<tr>
<td>Persons involved in or affected by the evaluation should be identified so their needs can be addressed.</td>
</tr>
<tr>
<td>B. Evaluator Credibility</td>
</tr>
<tr>
<td>Persons conducting the evaluation should be both trustworthy and competent to perform the evaluation so its findings achieve maximum credibility and acceptance.</td>
</tr>
<tr>
<td>C. Information Scope and Selection</td>
</tr>
<tr>
<td>Information collected should be broadly selected to address pertinent questions about the programme and be responsive to the needs and interests of clients and other specified stakeholders.</td>
</tr>
<tr>
<td>D. Values Identification</td>
</tr>
<tr>
<td>The perspectives, procedures, and rationale used to interpret the findings should be carefully described so the bases for value judgements are clear.</td>
</tr>
<tr>
<td>E. Report Clarity</td>
</tr>
<tr>
<td>Evaluation reports should clearly describe the programme being evaluated, including its context, purposes, procedures, and findings so that essential information is provided and easily understood.</td>
</tr>
<tr>
<td>F. Report Timeliness and Dissemination</td>
</tr>
<tr>
<td>Significant interim findings and evaluation reports should be disseminated to intended users so they can be used in a timely fashion.</td>
</tr>
<tr>
<td>G. Evaluation Impact</td>
</tr>
<tr>
<td>Evaluations should be planned, conducted, and reported in ways that encourage follow-through by stakeholders to increase the likelihood that the evaluation will be used.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feasibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Practical Procedures</td>
</tr>
<tr>
<td>The evaluation procedures should be practical to keep disruption to a minimum while needed information is obtained.</td>
</tr>
<tr>
<td>B. Political Viability</td>
</tr>
<tr>
<td>The evaluation should be planned and conducted with anticipation of the different positions of various interest groups so their co-operation may be obtained, and possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted.</td>
</tr>
<tr>
<td>C. Cost Effectiveness</td>
</tr>
<tr>
<td>The evaluation should be efficient and produce information of sufficient value so the resources expended can be justified.</td>
</tr>
</tbody>
</table>
Propriety

A. Service Orientation
Evaluations should be designed to help organisations address and effectively serve the needs of the full range of participants.

B. Formal Agreement
The obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to in writing to ensure that they adhere to all conditions of the agreement or that they formally renegotiate it.

C. Rights of Human Subjects
Evaluations should be designed and conducted to respect and protect the rights and welfare of human subjects.

D. Human Interactions
Evaluators should respect human dignity and worth in their interactions with other persons associated with an evaluation so participants are not threatened or harmed.

E. Complete and Fair Assessment
The evaluation should be complete and fair in its examination and recording of strengths and weaknesses of the programme being evaluated so that strengths can be built upon and problem areas addressed.

F. Disclosure of Findings
The formal parties to an evaluation should ensure that the full set of evaluation findings along with pertinent limitations are made accessible to the persons affected by the evaluation, and any others with expressed legal rights to receive the results.

G. Conflict of Interest
Conflict of interest should be dealt with openly and honestly so it does not compromise the evaluation processes and results.

H. Fiscal Responsibility
The evaluator's allocation and expenditure of resources should reflect sound accountability procedures, and otherwise be prudent and ethically responsible to ensure they are accounted for and appropriate.

Accuracy

A. Programme Documentation
The programme being evaluated should be described and documented clearly and accurately.

B. Context Analysis
The context of the programme should be examined in enough detail so its likely influences can be identified.

C. Described Purposes and Procedures
The purposes and procedures of the evaluation should be monitored and described in enough detail so they can be identified and assessed.

D. Defensible Information Sources
The sources of information used in a programme evaluation should be described in enough detail so their adequacy can be assessed.

E. Valid Information
The information-gathering procedures should be chosen or developed and implemented to ensure that the interpretation is valid for the intended use.

F. Reliable Information
The information-gathering procedures should be chosen or developed and implemented to ensure that the information is sufficiently reliable for the intended use.

G. Systematic Information
The information collected, processed, and reported in an evaluation should be systematically reviewed, and any errors found should be corrected.

H. Analysis of Quantitative Information
Quantitative information should be appropriately and systematically analysed so evaluation questions are effectively answered.

I. Analysis of Qualitative Information
Qualitative information should be appropriately and systematically analysed so evaluation questions are effectively answered.

J. Justified Conclusions
The conclusions reached in an evaluation should be explicitly justified so stakeholders can assess them.

K. Impartial Reporting
Reporting procedures should guard against distortion caused by personal feelings and biases of any party to the evaluation so that evaluation reports fairly reflect the evaluation findings.

L. Meta-evaluation
The evaluation itself should be formatively and summatively evaluated against these and other pertinent standards so that its conduct is appropriately guided, and, on completion, stakeholders can closely examine its strengths and weaknesses.
UNICEF Evaluation Report Standards

The UNICEF Evaluation Report Standards have been created as a transparent tool for quality assessment of evaluation reports. This document outlines what the Standards are, the rationale for each standard and how they are applied. The Standards are used by UNICEF Evaluation Office to assess evaluations for inclusion in the organisation’s Evaluation and Research Database to strengthen the Database as a learning tool. Application of Standards will also provide feedback to UNICEF Country Offices on how the evaluation is seen by someone outside of the evaluation process.

The Standards are also intended for use by UNICEF offices and partners commissioning evaluations to establish the criteria against which the final report will be assessed. The UNICEF Evaluation Report Standards draw from and are complementary to key references on standards in evaluation design and process increasingly adopted in the international evaluation community.

Evaluation Office
UNICEF NYHQ
September 2004
Table of Contents

I. Introduction and Procedures ................................................................................................... 1

II. Evaluation Report Standards

1. Completeness of Title Page and Opening Pages ................................................................. 6

2. Assessment of Executive Summary ..................................................................................... 6

3. The project/programme to be evaluated was clearly described, including the logic of the programme design or expected results chain. ........................................................................ 7

4. The role and contributions of UNICEF and other stakeholders to the project/programme were clearly described ........................................................................................................ 8

5. Purpose and context provided a specific justification for undertaking this evaluation at this time. ............................................................................................................................................. 9

6. The evaluation used standard OECD/DAC evaluation criteria as per UNICEF PPPM and/or provided an explanation for criteria that was considered not applicable and not used .............. 10

7. Objectives were realistic and achievable in light of the information that can be collected in the context of the programme/project. The scope of the evaluation was defined clearly ............. 11

8. The evaluation design considered programme/projects’ incorporation of the UN and UNICEF’s commitment to human rights-based approach to programming ................................................. 12

9. The evaluation design considered programme/projects’ incorporation of results based management ........................................................................................................................................ 15

10. Transparent description of methodology ........................................................................ 16

11. Evaluation methods were appropriate and adequate providing a complete and fair assessment. Consideration was given to limitations of the methodology ..................................................... 17

12. A complete description for stakeholder participation was given ........................................ 18

13. Where information was gathered from those who benefited from the programme/project, information was also gathered from eligible persons not reached ................................................................ 19

14. The evaluation design was ethical and included ethical safeguards where appropriate, including protection of the confidentiality, dignity, rights and welfare of human subjects, particularly children, and respect for the values of the beneficiary community .......................................................... 20

15. In presenting the findings, inputs, outputs, and, where possible, outcomes/impacts were measured (or an appropriate rationale given why not) ............................................................................. 21

16. To the extent feasible, the report includes cost analysis ..................................................... 22

17. Analysis of results includes appropriate discussion of the relative contributions of stakeholders to results ................................................................................................................................. 23
18. Reasons for accomplishments and difficulties, especially continuing constraints, were identified as much as possible....................................................................................................... 24

19. Conclusions were substantiated by findings consistent with data and methods and represent insights into identification and/or solutions of important problems or issues. ............................... 25

20. Recommendations were firmly based on evidence and analysis; they were directly relevant and realistic with priorities for action made clear........................................................................... 25

21. Lessons learned, when presented, were generalized beyond the immediate intervention being evaluated to indicate what wider relevance there might be. ......................................................... 26

22. Completeness of Annexes...................................................................................................... 27

III. Glossary of Terms.................................................................................................................. 29
I. Introduction and Procedures

The UNICEF Evaluation Report Standards have been created as a transparent tool for quality assessment of evaluation reports. This introduction provides an overview of why and how the Standards were developed and how they are used. The body of this document outlines, for each of the 22 Standards, what is being assessed, why this is important and how it is applied.

The Standards are intended to be useful to both UNICEF and partners in the management of programme evaluation. These Standards draw from and are complementary to a broad range of widely accepted references on evaluation design and process. They also set standards for how evaluation reports address key shared approaches in development and humanitarian practice – specifically results based management and a human rights based approach to programming, both of which figure explicitly as guiding principles for how UNICEF works.

WHY the Standards were developed

In the past, all evaluations have been provided to UNICEF staff through EvalWin. An analysis of EvalWin and meta-evaluations of UNICEF’s evaluations have found that the uneven quality of evaluations has limited their use as a learning tool. Staff are more interested in having good reports rather than having all reports. The Standards will be used to determine which evaluations are posted on the Evaluation and Research Database on the Intranet and Internet.

The Standards also provide feedback on how the evaluation is seen by someone outside of the evaluation process. The office will have another perspective on where the report needs additional clarity or explanation allowing offices to gauge progress in evaluation report quality. This feedback can be seen by offices on the Evaluation Report Submission Website.

Additionally, the Standards are intended for use by UNICEF offices and partners commissioning evaluations to establish the criteria against which the final report will be assessed. For future meta-evaluations as well, Country Offices now know the standard by which the evaluation reports will be judged.

HOW the Standards were developed

First overall programme evaluation quality standards were consulted from:

- American Evaluation Association
- African Evaluation Association Guidelines
- Swiss Evaluation Society Standards


2 Overall evaluation quality standards developed by reputable international evaluation associations relate to the entire evaluation process from planning to the final report. The evaluation report standards outlined in this document are complementary to overall evaluation quality standards. UNICEF offices are encouraged to know and use the overall evaluation quality standards developed by their area evaluation association. Given that there is a large degree of overlap of these guidelines, it is reasonable to expect that the UNICEF evaluation report standards will not pose any conflict to overall evaluation quality standards from various area evaluation associations. Areas of significant disagreement should be brought to the attention of the Evaluation Office (EO).

3 http://www.eval.org/EvaluationDocuments/progeval.html
4 http://www.afrea.org/content/index.cfm?navID=5&ItemID=204
Evaluation Report Standards

- OECD/DAC Principles for Evaluation\(^6\)

and standards for evaluation reports were compiled from:
- meta-evaluations mentioned above
- ALNAP Quality Proforma\(^7\)
- UNICEF M&E Training Resource

Drawing on the above, draft evaluation report standards were created. The following factors were considered in formulating these Standards:
- trends across evaluation standards
- usability given current evaluation capacity and a manageable set of standards
- UNICEF specific issues such as Human Rights Based Approach to programming

Feedback was then gathered from M&E officers and focal points and other divisions.

WHAT the Standards are

For each standard the following information is given:

WHAT is the standard
WHY is this standard included
HOW is this standard applied
LIMITATIONS to the use of this standard
TIPS for report preparation

Throughout the report, the word “evaluator” refers to those involved in conducting the evaluation and preparing the evaluation report. “User” refers to anyone who might read the evaluation report for any purpose. “Reader” refers to those UNICEF staff who will be applying these Standards to the evaluation report.

Words that are underlined in blue are linked to a glossary at the end of the document. Holding down the Cntrl key on the keyboard while clicking on the word with the mouse will bring the cursor to the definition in the glossary.

It is planned to issue another version of these Standards after they have been in use for six months. The new version will include examples of many items based on the experience gained during that time period. Throughout the report, the EO will add “TIPS for report preparation” that will link to the UNICEF M&E Training Resource once the updated manual is finalized.

Responsibilities

As stated in the most current UNICEF Programme and Policy Procedures Manual and audit guidelines, it is the responsibility of the UNICEF commissioning office to include these Evaluation Report Standards in every TOR for an evaluation and explicitly mention these Standards in describing the outputs of the TOR. It is also thus assumed that the Standards will be referred to when reviewing the end product to ensure that it complies with the TOR. Beyond this, there is no required procedure, no formal consequence from UNICEF NYHQ’s point of view. A write-up on

\(^6\) http://www.oecd.org/findDocument/0,2350,en_2649_34435_1_119820_1_1_1,00.html
\(^7\) ALNAP Annual Review, 2003; http://www.alnap.org

23/9/2004
the application of these Standards is not required by NYHQ. The UNICEF commissioning office or RO can introduce further uses as suitable to their situation.

Disclosure Policy
For the Evaluation and Research Database on the Intranet, all evaluations that meet these Standards will be posted. Evaluations that meet these Standards will also be posted on the parallel Evaluation and Research Database on the Internet unless they are identified as being sensitive in nature by the UNICEF office that commissioned the report. As established by the Evaluation Committee Rules and Procedures, June 2004, the EO will assume that all reports are suitable for public dissemination unless informed in writing by the commissioning office. Further information will be provided in a forthcoming Executive Directive. Partners who wish an evaluation report to remain internal are asked to contact the UNICEF office that participated in the evaluation with them and not to contact UNICEF NYHQ directly.

Application of Standards
The Evaluation and Research Database Editor, or a suitable consultant, will rate every evaluation report. The Evaluation Officer supervising the ERD will perform random checks on the ratings to ensure consistency and accuracy. If an office disagrees with the rating given, they are encouraged to contact the UNICEF Evaluation Office (EO). The EO welcomes all suggestions and will work with the office until a common understanding is reached. (As explained below in the Dissemination of Ratings section, only “satisfactory”, “very good” or “excellent” evaluations will be seen by everyone. “Poor” evaluations are only seen by the UNICEF CO that commissioned the evaluation, and by the ROs and HQ.)

Evaluations and their rating will be regularly checked with a staff member in the appropriate Division or technical cluster to confirm the quality of the report for those evaluations that the EO would like to either: profile on the Intranet (write up a short “Lessons from Evaluation” or similar article); or use as a methodological example for that sector.

There are a variety of options for an office to facilitate a good quality review and consistency with the Standards during the evaluation cycle. For example:

- Evaluation team consults the ERD; the person who will be authoring the evaluation report may want to read some reports that are rated “Excellent” in the ERD;
- Evaluation manager checks through the evaluation Standards in designing the TOR;
- Evaluation team reads through the evaluation Standards while organizing their data and formulating the report outline; and/or
- Evaluation manager and author together read through the Standards after the first draft has been finished but before the final draft.

There are some options for the office after the report has been released:

- If the consultant is still available, the office can work with the consultant and relevant stakeholders to adjust the report – adding description or clarifying sections – and issue a new version. Clearly, findings and conclusions cannot be changed to any significant extent without the original author and relevant stakeholders’ agreement.
- If the consultant is not available and the office finds the evaluation weak enough as to make the recommendations unreliable, an office may always write a formal response to an evaluation describing what recommendations they accept and their action plans based on

---

8 The one exception is evaluations in non-UN languages. In this case, COs are to apply the Standards and provide a write-up. A rate on each standard with full comments (in English) and an English Executive Summary should be sent to the Evaluation Office via email to Elizabeth Santucci.
those recommendations; and what recommendations they do not accept and why. As of 2004, Country Offices are required to establish their response to recommendations through the UNICEF Country Management Team or another suitable body as appropriate to which office(s) the recommendations are addressed.

If an office decides to apply the Standards to an evaluation report themselves and finds that it is below “satisfactory” and is not adjusting the evaluation, the evaluation still must be sent to the EO for institutional memory purposes.

Relation to Institutional Memory Databases

The Evaluation Office (EO) has the responsibility to maintain the organizational memory of evaluation in UNICEF. All evaluations – conducted or commissioned by the COs, the ROs, and headquarters – are stored in an archival database. Country Offices should send all evaluations (and recommended studies and surveys) to the EO using the Evaluation Report Submission Website accessible on the UNICEF intranet. These reports are registered in the system and the electronic copies of the reports are transferred to a secure electronic archival system called Ramp-Trim.

UNICEF staff still have access to all evaluations upon request. The EO can conduct searches of the Archive for staff members and provide evaluations that are not in the ERD on the Intranet. Additionally, the EO can provide a list of evaluations conducted by the COs that the EO has not yet received for the staff member to contact the CO directly him/herself.

Maintaining this institutional memory is a critical component of the EO and the office’s accountability. Evaluation is an important part of programme learning and improvement that should be available to all.

Rating

For each standard in the section “HOW is this standard applied,” an explicit description specific to that standard of what constitutes “poor”, “satisfactory”, “very good” and “excellent” is given. There are also basic quality factors that the reader will consider in determining a rank. For each standard, the applicable section of the report will also be assessed for:

- Length – not too long
- Coverage – comprehensiveness
- Linkage – relevant to overall purpose of report and consistent with other elements of report
- Clarity – clear and simple language, understandable to outside reader

If any of these qualities is felt to be lacking, the report will not receive “excellent” for that standard.

The average score needed to merit inclusion in ERD is “satisfactory”, “very good” or “excellent.” In general, “excellent” represents a model for all evaluations in this area and will be infrequently used. The standards will be weighted – the mark on each standard will count differently towards the average based on the weight given the standard.

In addition to having an average score of “satisfactory” or better, the report must also receive “satisfactory” or better on each of the four key standards below. If the report receives an average score of “satisfactory” or better but receives less than that on any of the key standards, it will not be included in the ERD. The proposed key standards are:

10. Transparent description of methodology
11. Evaluation methods were appropriate and adequate providing a complete and fair assessment. Consideration was given to the limitations of the methodology.

*Reports completed prior to implementation of these Standards that fail either #10 or #11* - will be considered acceptable for inclusion in the ERD if the evaluation receives an average score of “satisfactory” or greater.

*Reports completed during the implementation period from September 2004 to April 2005 that fail either #10 or #11* – if the report receives a passing average score but does not pass either of these standards, an email will be sent to the office that commissioned the report pointing out the importance of clear and sound methodology that includes an awareness of the limitations. The report will be considered acceptable and included in the ERD.

*Reports completed after April 2005 that fail either #10 or #11* - the report will not be included in the ERD regardless of the overall rating.

14. The evaluation design was ethical and included ethical safeguards where appropriate, including protection of the confidentiality, dignity, rights and welfare of human subjects, particularly children, and respect for the values of the beneficiary community.

*Reports completed prior to implementation that fail #14* - will be considered acceptable for inclusion in the ERD if the evaluation appears to the reader to be ethically designed and conducted and if the evaluation receives an average score of “satisfactory” or greater.

*Reports completed during the implementation period from September 2004 to April 2005 that fail #14* – if the report receives a passing average score but does not pass this standard, an email will be sent to the office that commissioned the report pointing out the importance of this issue and the need to include a description of the ethical safeguards used during the evaluation in the evaluation report. The report will be considered acceptable and included in the ERD.

*Reports completed after April 2005* - if no statement concerning ethical safeguards is provided, the report will not be included in the ERD regardless of the overall rating.

15. Conclusions were substantiated by findings consistent with data and methods and represent insights into identification and/or solutions of important problems or issues.

This key standard will be implemented directly. Any evaluation – even for those evaluations already completed and received by NYHQ – that does not pass this standard will not be included in the ERD.

**LIMITATIONS to the use of the Standards**

These Standards are applicable to the evaluation report. The quality of any evaluation is truly measured by its planning and implementation – the actual report can only hold a mirror to this. Although the Standards indicate what a good evaluation is, the Standards are no replacement for the M&E Training Resource and other manuals which should be consulted and used.
II. Evaluation Report Standards

1. Completeness of Title Page and Opening Pages

WHAT is the standard

The following information should be easily accessible in the first few pages of the report:

- Name of programme or project being evaluated
- Date
- Country/ies of programme/project
- Name and organizations of the evaluators
- Name and address of organization(s) the report has been commissioned by
- Name of UNICEF staff contact point for the evaluation (someone involved in the evaluation not someone responsible for the evaluation files)
- Table of contents
  - with the contents of the Annexes listed

WHY is this standard included

This basic information is needed to keep track of the report if it were to be separated from a cover letter or surrounding materials. There have been cases where the name of the country does not appear anywhere in the report and it must be deduced from the city names of the sample!

HOW is this standard applied

A simple check list is used to mark off that each of these items is there.

2. Assessment of Executive Summary

WHAT is the standard

An Executive Summary should provide an overview of the essential parts of a report. It should be very short – ideally two to three pages – and should “stand alone” (without requiring reference to the rest of the report.) The Executive Summary should include:

- Brief description of programme/project
- Context of programme/project – years of implementation, situation vis-à-vis UNICEF Country Programme outcomes and other programming it contributes to (i.e. UNDAF outcomes, complementary national or partner programmes)
- Basic description of context and purpose of evaluation – why this evaluation now
- Objectives of evaluation
- Short description of methodology; key features include:
  - Rationale for choice of methodology
  - Data sources used
  - Data collection and analysis methods used
  - Major limitations
- Most important findings and conclusions
- General areas of recommendation
  - With highest priority recommendations given in more detail

WHY is this standard included
Evaluation Report Standards

Primarily for key decision-makers who do not have time to read the full evaluation report, a good Executive Summary increases the utility of the evaluation. Additionally, the UNICEF Evaluation and Research Database (ERD) now uses the full Executive Summary for each report.

HOW is this standard applied

A four point scale is used:
- Missing – the report lacks an Executive Summary
- Poor – the Executive Summary does not provide the whole picture, leaving out essential information, either the name of programme, findings or recommendations
- Satisfactory – the Executive Summary provides a clear picture but does not encompass all of the elements above (perhaps missing a critical finding or recommendation or an aspect of the methodology) or cannot stand-alone from the report
- Very good - the report has an Executive Summary that includes all of the elements above and can effectively stand-alone from the report
- Excellent – a model for this kind; clear, concise and could be used to inform decision making

LIMITATIONS to the use of this standard

It may be necessary to draft different versions of Executive Summary for different audiences. Not all versions will include all of the elements recommended in this standard. The content described above is, however, necessary for outside users and must be included in the version provided for the Evaluation and Research Database.

TIPS for report preparation

Standards for an executive summary are detailed in Technical Note #3, “Writing a good Executive Summary” available on the Intranet. An Executive Summary should be included in all evaluations and the Terms of Reference should oblige the evaluator to provide one. The ERD can also be used as a source for examples of good executive summaries. (For reports posted on the ERD, the EO will alter the Executive Summary to meet the standard above if necessary.)

3. The programme/project to be evaluated was clearly described, including the logic of the programme design and/or expected results chain.

WHAT is the standard

The user not only needs to know what the programme/projects does, but how the designers of the programme/project thought the programme/project would address the problem they had identified. The overall goal that the programme/project is expected to contribute to and how it was expected to contribute should be described. The overall goal may be related to national development plans or MDGs which may in turn be supported by UNICEF’s MTSP, MDG or WFFC goals. Critical elements of the national or sub-national context which have shaped the programme design needs to be explained.

Additionally, how the programme was intended to be implemented needs to be explained if the evaluation finds/documents a change in programme implementation (intentional or unintentional change).

http://www.intranet.unicef.org/epp/evalsite.nsf/1565f9b3780158a285256b95005a5231/81db8238743b9ed985256c31004ef905?OpenDocument
Other important elements are the size of the programme/project – including the relative size of the programme/project to the Country Programme. At the very least, description should include the number of participants/people reached by the programme/project. Budget figures are important though there are occasions where this information is hard to obtain.

**WHY is this standard included**

A good description of the project is essential to increase the utility of the evaluation to other users in addition to providing general clarity. Users not familiar with UNICEF’s work (or UNICEF staff involved in other sectors/geographic areas) should clearly understand what the project does. This allows them to draw parallels to their own programme/project and see how they can apply findings and recommendations to their own situation.

Users and evaluators alike must have an understanding of the programme logic to accurately situate the programme results and the effects of the wider context.

Even if a report has an Executive Summary, these background materials are still necessary. An Executive Summary does not take the place of an Introduction – they serve different purposes with different components.

**HOW is this standard applied**

- Missing
- Poor – a vague programme/project description is given that does not paint a complete picture for the user
- Satisfactory – the programme/project is described but gaps in logic or relevant context are unexplained
- Very good - the programme/project is well described including its logic model and relevant context
- Excellent – a model for this standard providing a concise overview that illuminates the findings and analysis of outcomes

**LIMITATIONS to the use of this standard**

The programme logic and expected results chain may not be well described in the programme documentation. This is not a problem in the evaluation but in programme design. A good evaluation, however, will note this and may have developed a programme logic retroactively with key programme staff. If this is done, care should be taken that the programme logic built retroactively should focus on the intended implementation and outcomes and document any adjustments to the implementation. The observed implementation and outcomes – or findings – should not be assumed to be the same as what was planned.

**TIPS for report preparation**

This standard should not encourage a lengthy description of the programme/project. If additional details of the history or logic model are pertinent to the findings and conclusions, a full description and logic model can be provided in an annex. The evaluator should ensure that the pertinent information is given without being excessively long.

**4. The role and contributions of UNICEF and other stakeholders to the programme/project were clearly described.**

**WHAT is the standard**

Describe who is involved, in what role and what they have contributed to the programme/project including: financial resources, in-kind contributions (material items such as drugs, books, desks,
etc.), technical assistance, participation, staff time, training, leadership, advocacy and lobbying. This should include any contributions from primary stakeholders, which includes children and adolescents. The reader needs to have a sense of who is doing what that facilitates an understanding of what partners assisted in which outcome.

**WHY is this standard included**

A statement as to who was involved in the programme and in what role is important background information. Users will want to compare this with who was involved in the evaluation to assess how different points of view were included. Users will also need to refer back to this in understanding evaluation findings and outcomes.

**HOW is this standard applied**

- Missing
- Poor – only a single party’s contributions are identified OR the information is scattered throughout the report without any brief summary in the introduction
- Satisfactory – all stakeholders contributions are identified but not a complete picture such as only in a single dimension, i.e. just financial inputs are described with no attention to staff resources, community members’ time and labour, etc.
- Very good – all stakeholders’ contributions are comprehensively described in an easily understandable manner that aids understanding of the outcomes
- Excellent - a model for this standard, would include what primary stakeholders may have contributed to the programme/project and what type of support UNICEF provided

**TIPS for report preparation**

The reader should not have to search all over the report for this information. The programme description should be a complete overview in one section. If the reader sees primary stakeholder involvement through school construction in the findings, for example, but did not read about such involvement in the programme description, then that description was lacking and could not be considered “Excellent” for this standard.

5. **Purpose and context is described providing a specific justification for undertaking this evaluation at this time.**

**WHAT is the standard**

Purpose discusses why the evaluation is being done and how it will be used. It should not be confused with the evaluation objectives - which state what the evaluation seeks to accomplish. The purpose also relates to the timing of the evaluation in the project cycle - at the beginning of the project, mid-way through the project, after the project is completed. Sometimes it may be relevant to relate the project to the Country Programme cycle, especially if the evaluation is to contribute to a Mid-Term Review.

The reader is basically looking for why, why now, and how will the evaluation be used. Explaining why the evaluation is being done - what triggered the evaluation - does not always answer how it will be used by the stakeholders – who will use it, what decisions will be taken after the evaluation is complete. For example, a report may state that an evaluation is being done at the end of a five-year programme. The reader is left wondering if the evaluation is to assist the programme for its next cycle or to determine how to hand the programme over to the government or another NGO or for general lessons learned for similar programmes now that this particular programme has finished. Evaluations are most used when they are planned to coincide or are driven by a decision that needs to be taken. Ideally, an evaluation report is also intended to have use value to partners and other stakeholders.
Evaluation Report Standards

WHY is this standard included

This adds to the clarity of the evaluation and should provide the broad orientation which is then further elaborated in the objectives and scope of the evaluation – why something is evaluated sets up more clearly what the evaluation should do. Purpose and timing should help to define what questions the evaluation should answer.

HOW is this standard applied

- Missing
- Poor – Purpose is unclear or does not correlate with objectives and scope.
- Satisfactory – Purpose is clear but context of evaluation is either missing or unclear
- Very good – Purpose and context are clearly described and relates appropriately to objectives and scope
- Excellent - a model for this standard

TIPS for report preparation

For some evaluations, the purpose may seem self evident or the evaluation may have to conform to a model format for a particular client – such as an evaluation at the end of the programme/project cycle for a donor report. However, to those removed from the context, the purpose is not apparent. The evaluation report should always explicitly state the purpose, context and objectives in the report.

6. The evaluation used standard OECD/DAC evaluation criteria as per UNICEF PPPM and/or provided an explanation for criteria that was considered not applicable and not used.

WHAT is the standard

UNICEF manuals refer to the following established OECD/DAC evaluation criteria:

- Relevance
- Efficiency
- Effectiveness
- Impact
- Sustainability

And for evaluations of humanitarian response should also include:

- Coverage
- Co-ordination
- Coherence
- Protection

Not all criteria are applicable to every evaluation. Which criteria to use is determined by the evaluation’s objectives and scope. The rationales for not using a particular criterion should be explained in the report.

Performance standards or benchmarks should be identified. For example, statements that qualify the programme outputs such as “acceptable range” should have some clear reference standard. Where appropriate, international standards should be referred to in defining benchmark terms.
Any foreseen limitations in applying the evaluation criteria should also be noted. This may be presented in a broader discussion of evaluation questions and scope; both of which are also covered under standard #9.

**WHY is this standard included**

It is imperative to make the basis of the value judgments transparent if the report is to be understood, considered convincing, and accepted by stakeholders. Additionally, the established OECD/DAC criteria was developed through common consensus of experienced evaluation offices and organizations. By considering each criteria, it is ensures that basic evaluation question are covered and decisions to limit evaluation scope through dropping one or more of these criteria are conscious and considered.

**HOW is this standard applied**

- **Missing** - do not use OECD/DAC criteria or objectives could only be considered OECD/DAC with broad interpretation by the reader
- **Poor** - objectives easily translate to OECD/DAC criteria, but no explanation as to why all OECD/DAC criteria was not used OR criteria selected is not appropriate to the purpose of the evaluation
- **Satisfactory** - some OECD/DAC criteria used and applicable but no explanation as to why all OECD/DAC criteria was not used
- **Very good** - OECD/DAC criteria used, applicable with a description of why all OECD/DAC criteria was not used OR different criteria are well described, applicable and a sound explanation given as to why OECD/DAC criteria was not used
- **Excellent** – a model for this standard that explains how the criteria will be used to provide a comprehensive assessment of the programme, linking the evaluation purpose and context to the evaluation objectives and scope

**LIMITATIONS to the use of this standard**

An evaluation can have clear objectives, be methodologically sound, and be relevant but miss key questions – for example, an evaluation that looks only at effectiveness and ignores relevance and sustainability. The report in this case would score low on this standard while still having high marks for objectives and methodology. It represents a lost opportunity to find out more about how to improve programming.

On the other hand, there are evaluation reports that list only one or a few criteria in the methodology section yet provide analysis using additional or alternate criteria in the findings. The report in this case could not be considered Very Satisfactory or Excellent.

Evaluation criteria is usually given in the TOR for the evaluation. The TOR may be drafted and criteria selected without involving the person(s) commissioned to conduct the evaluation. In this situation, the evaluator should work with the evaluation manager to adjust the criteria for analysis appropriately.

**TIPS for report preparation**

There are resources available that further define and explain these terms. Please see the UNICEF intranet and/or the PPPM.

7. **Evaluation objectives were realistic and achievable in light of the information that can be collected in the context of the programme/project. The scope of the evaluation was defined clearly.**
WHAT is the standard

The objectives of the evaluation should follow from the purpose for the evaluation outlined above. They should be clear to all individuals and groups involved in the evaluation. Ideally, they should be SMART - Specific, Measurable, Achievable, Relevant, and Time-bound.

Scope determines the boundaries of the evaluation, tailoring the objectives and evaluation criteria to the given situation. It should also make the coverage of the evaluation explicit (time period, phase in implementation, geographic area, and the dimensions of the network of actors being examined, i.e. a single organisation or all stakeholders.) The limits of the evaluation should also be acknowledged within the scope.

Evaluations may also be oriented by evaluation questions. These add more detail to the objectives and contribute to defining the scope.

WHY is this standard included

This is the critical reference against which the findings/conclusions and recommendations are compared. The objectives and scope of the evaluation are also the critical reference to judge whether the methodology selected and resources allocated were adequate.

HOW is this standard applied

- Missing
- Poor – objectives are vague, not related to the purpose or impractical given the context
- Satisfactory – objectives are clear but: either the scope is not clear OR it is not entirely clear how the objectives will be measured and achieved leaving the evaluator to feel their way
- Very good – objectives and scope are clear, measurable, achievable and relevant
- Excellent – a model for this standard, objectives were clear, measurable, achievable, relevant and related to the evaluation purpose and OECD/DAC criteria as in Standard #6

LIMITATIONS to the use of this standard

A thread of logic should pass from the description of the programme/project and the programme/project objectives, to the purpose of the evaluation, to the objectives and scope of the evaluation. Each may be clear and solid on its own but they may not be clearly related to one another. The score and comments on the Standards together should reflect this thread and not fall in the trap of treating each piece in isolation.

An evaluation may have unrealistic or vague objectives yet collected good information with a sound methodology and supported findings. If this occurs, the score on this standard would be quite low while the report can still receive high scores on methodology and findings.

The TOR may be drafted without involving the person(s) commissioned to conduct the evaluation and have poor objectives. In this case, the evaluation manager and evaluator should work together with other relevant stakeholders to reframe the objectives. An excellent evaluation report would describe why the original objectives were un-workable and what changes were made to the evaluation design.

8. The evaluation design considered programme/projects’ incorporation of the UN and UNICEF’s commitment to human rights-based approach to programming.
WHAT is the standard

Children, as rights-holders, have claims against those with obligations to ensure the fulfilment of those rights. Children whose rights remain unmet have claims against those with an obligation to act. Parents, communities, civil society organizations, governments, and others have resulting duties. At the same time, parents and other duty bearers also may have unfulfilled rights, for example due to poverty. Vulnerability and exclusion are manifestations and causes of the lack of capacities within families, communities, government, and others to fulfil children's rights.10

A Statement of Common Understanding was developed in the Interagency Workshop on a Human Rights based Approach in the context of UN reform 3-5 May, 2003 that identifies the necessary elements for a HRBAP:

. . . the application of “good programming practices” does not by itself constitute a human rights-based approach, and requires additional elements.

The following elements are necessary, specific, and unique to a human rights-based approach:

a) Assessment and analysis identify the human rights claims of rights-holders and the corresponding human rights obligations of duty-bearers as well as the immediate, underlying, and structural causes of the non-realization of rights.

b) Programmes assess the capacity of rights-holders to claim their rights, and of duty-bearers to fulfil their obligations. They then develop strategies to build these capacities.

c) Programmes monitor and evaluate both outcomes and processes guided by human rights standards and principles.

d) Programming is informed by the recommendations of international human rights bodies and mechanisms.

These four elements rather neatly encapsulate much of the material developed at UNICEF regarding HRBAP. In programming, UNICEF must be mindful of the articles of the CRC and CEDAW and of the guiding principles of these and other human rights treaties.

Within this analysis, a gender perspective is critical, particularly in regard to understanding the often more disadvantaged status of women and girls. “Children experience discrimination in various other dimensions in addition to gender, such as ethnicity, language, disability and rural-urban residence. Unfortunately, when a child is female, this usually places her at a double disadvantage.”11 For UNICEF, gender issues have been implemented as a cross-cutting theme in programming and the organization has given promoting gender equality and gender-sensitive development programmes high priority.12 As such, a description of the consideration of gender issues is a minimum requirement for all evaluation reports.

At this point in time, UNICEF does not have a uniform methodology to assess the implementation of the Human Rights Based Approach to Programming. It is necessary to develop an appropriate assessment framework depending on the programme/project and context. In all cases a gender sensitive perspective should be incorporated.


11 MTSP, para 40, p. 17

The evaluation report should describe how the programme/project:
− was informed by the CRC and CEDAW;
− identified the human rights claims and obligations relevant to the programme;
− identified gaps in the capacity of rights-holders to claim their rights, and of duty-bearers to
  fulfil their obligations, including (but not solely) an analysis of gender and marginalized and
  vulnerable groups;
− how the design and implementation of the programme addresses these gaps; and
− how the programme monitored results and viewed results within this rights framework.

WHY is this standard included

UNICEF, as part of the United Nations system and guided by the United Nations Charter, has a
responsibility and mission to help State parties to human rights treaty bodies to meet their
obligations towards the realisation of the human rights of those who live within their jurisdiction.
As the UNICEF Mission Statement makes clear, the Convention on the Rights of the Child is the
organisation’s guiding frame of reference, and provides a legal foundation for the ethical and
moral principles that guide the work of UNICEF for children. The other keystone of the
organisation’s mandate and mission is the Convention on the Elimination of All Forms of
Discrimination Against Women. Both the CRC and the CEDAW comprehensively address the
social, economic, cultural, civil and political rights of children and women. The scope of UNICEF
involvement therefore includes areas of concern with any or all of these categories of rights.13

A consideration of HRBAP is essential to ensuring the organization’s commitment to these
principles. As such, it is part of the standard criteria above concerning relevance – specifically
policy relevance. It is highlighted separately here because of the overall importance of this
guiding principle for UNICEF.

Additionally, attention to programme effects on marginalised, vulnerable and hard to reach groups
is necessary to gain a complete and accurate assessment. This standard also reinforces credible
methodology. HRBAP is emerging as a standard of good programming; as such an analysis
based on human-rights considerations is increasingly considered a prerequisite for a credible
evaluation.

HOW is this standard applied

• Missing – no description of the programme’s HRBAP or of any gender analysis
• Poor – minimal consideration given to this strategy, i.e. a token paragraph considering
  programmes link to the CRC or CEDAW but not with a specific analysis of the rights
  claims and duties for this programme area
• Satisfactory – report considers human rights with a systematic description of the
  programme/projects consideration of rights claims and duties (including an analysis of
  gender) but not in a methodological way; i.e. the description is limited to a one time
  statement concerning programme design and is not used in analysis of all phases of
  programming - missing in programmes monitoring and results
• Very good – an analysis of HRBAP is incorporated into the evaluation framework in an
  appropriate and context specific manner, including a gender sensitive perspective, with
  analysis applied to all phases of programming
• Excellent – a model for this standard that UNICEF can use to develop standard
  methodology on this topic

13 PPPM, 2004, Chapter 2,
http://www.intranet.unicef.org/Policies/DHR.nsf/cc58cfbb4d01337f85256720005e2cd7/77a963b4ce562e37985256bd7006ccce9?OpenDocument
LIMITATIONS to the use of this standard

Given the lack of Standards for evaluation in this area, evaluators will be feeling their way. It is understandable that it may be somewhat imprecise. A variety of sources are available to UNICEF staff on the Intranet to assist them in framing the evaluation and to provide guidance to the evaluators in this area. As our work in this area progresses, further guidance will be given and more detail provided in subsequent versions of these Standards.

This standard – especially when considered together with the next standard - provides a comprehensive view of the overall programme design. A reminder – this is not the standard by which to judge the programme itself but rather the evaluation’s analysis of the programme. An evaluation report can receive an “Excellent” score even if the programme was poorly designed.

9. The evaluation design considered programme/projects’ incorporation of results based management.

WHAT is the standard

Results Based Management is a management approach “focusing on performance and the achievement of outcomes and impacts.”14 In examining relevance, efficiency, effectiveness, outcomes/impact and sustainability of a programme, an evaluation should explore some underlying questions as to how a programme is managed and what information stakeholders at different levels have had access to and used to take key decisions in design and implementation – such as what indicators the programme/project had planned to use to monitor programme/project performance, what indicators were used, how they were used, and how they influenced adjustments to the programme/project. These indicators may not be explicit – an excellent evaluation documents the reasons for programme adjustment working with programme staff to understand why they decided adjustments needed to be made and comments on whether this process was driven by a focus on results for children and women.

“The standard of programme excellence sought in the MTSP combines result-based management (RBM) with a human rights-based approach to programming (HRBAP).”15 Results in a human rights perspective must refer to both the achievement of desirable outcomes and the establishment of morally acceptable processes to achieve these outcomes. Results-based management therefore becomes a tool or a means to realize human rights.

WHY is this standard included

RBM is widely believed to be central to a programme achieving its desired outcomes and impact. It is highlighted here because of the overall importance of RBM in UNICEF corporate strategy.

HOW is this standard applied

- Missing – report does not comment on implementation or programme monitoring
- Poor – minimal consideration given to this strategy, i.e. a token paragraph speaking of results based management OR minor aspects – disparate tools - of RBM are commented on not as an overall management system, e.g. analysis of monitoring forms without a broader analysis of how managers managed results and adjusted the programme for results

14 OECD/DAC (2001) “Glossary of Evaluation and Results Based Management Terms”
Evaluation Report Standards

- Satisfactory – report considers the programme/project’s use of key elements of RBM such as results based planning or monitoring systems but not in a methodological way throughout the report, e.g. considers only the planning of results based indicators but not programme adjustment for results based upon these indicators
- Very good – an analysis of programme/project’s use of RBM is incorporated into the evaluation framework in an appropriate and context specific manner that is carried throughout the evaluation report
- Excellent – a model for this standard which would involve RBM fully integrated into the evaluation report in analysis, conclusions and recommendations

LIMITATIONS to the use of this standard

Explicit RBM has been increasingly introduced in UNICEF programming guidance over recent years. The recent DPP/EO guidance (September 2003) “Understanding Results Based Programme Planning and Management: Tools to reinforce good programming practice” sets out very clearly how long standing programming tools can be used to strengthen RBM. COs and partners may use alternative tools that have a strong RBM approach. Regardless of the tools used, excellent evaluations will ask simple clear questions exploring how programme management was results-focused in its design and implementation.

This standard – especially when considered together with the last standard - provides a comprehensive view of the overall programme design. A reminder – this is not the standard by which to judge the programme itself but rather the evaluation’s analysis of the programme. An evaluation report can receive an “Excellent” score even if the programme was poorly designed.

10. Transparent description of methodology

WHAT is the standard

This standard is to ensure that the critical aspects of the methodology are described fully in the report. The appropriateness and merit of the methodology is discussed in another standard below.

A description of the methodology should include:

- Data sources
- Description of data collection methods and analysis (including level of precision required for quantitative methods, value scales or coding used for qualitative analysis)
- Description of sampling – area and population to be represented, rational for selection, mechanics of selection, numbers selected out of potential subjects, limitations to sample
- Reference indicators and benchmarks, where relevant (previous indicators, national statistics, etc.)
- Key limitations

The appendices should include the following in addition to more detail on any of the above:

- Data collection instruments (surveys, checklists, etc.)
- System for ensuring data quality through monitoring of data collection and oversight
- A more detailed discussion of limitations as needed

WHY is this standard included

A full description of the methodology allows the user to come to their own conclusions about the quality of the data. Utility is increased by this assurance of quality. A clear description of
methodology contributes to the overall value of the report. Propriety is also ensured through transparency.

**HOW is this standard applied**

- Missing
- Poor – incomplete description is given lacking any of the following:
  - data sources
  - description of data collection methods
  - general description of data analysis
  - general description of sampling – population to be represented, rational for selection of sample, and methods of sample selection
  - key limitations
- Satisfactory – full description is given (see list under WHAT) but lacks a wider comparison - to international Standards, to past evaluations/indicators concerning this programme/project, or to secondary source materials
- Very good – full description of the methodology is given including a wider comparison
- Excellent – a model for this standard, comprehensive but not excessive

**LIMITATIONS to the use of this standard**

Full description of the methodology is sometimes presented in the Annexes to increase the readability of the report itself. Before determining a score on this standard, check the Annexes and supplemental materials.

**11. Evaluation methods were appropriate and adequate providing a complete and fair assessment. Consideration was given to limitations of the methodology.**

**WHAT is the standard**

The standard covers three elements:

- **Logical link to evaluation objectives** –
  The methodology should be suitable to the subject matter and the information collected should be sufficient to meet the evaluation objectives.

- **Methodology illustrates good practice** -
  Evaluation methods depend on the information sought and the type of data being analyzed. A complete description cannot be provided here.
  The data should come from a variety of sources to ensure its accuracy and also to ensure that all affected people/stakeholders are considered. Methodology should explicitly address issues of gender and under-represented groups.
  Commonly accepted practice for the given situation by evaluation professionals may also be considered and referenced.

- **Efforts to control bias and acknowledgement of limitations** –
  Limitations can come from a variety of sources both internal and external. Bias can be from three levels:
− Sources of data - the respondents themselves have a bias in their opinion on the topic
− Methods of data collection - the structure of the data gathering could be skewed to favour one factor, preconceived idea or viewpoint
− Analysis of data - the evaluators have a bias towards a certain viewpoint that colours their interpretation of the findings

Satisfactory methodology seeks to limit bias in design and to explicitly identify areas where bias may occur. Bias can be addressed through having a representative sample that seeks many different points of views to balance. The standard here should ensure that the choice of methodology and actions to limit bias provided a complete and fair assessment.

WHY is this standard included

This standard addresses basic issues of propriety and accuracy.

HOW is this standard applied

• Missing
• Poor – methodology is not suited to the subject matter or objectives
• Satisfactory - methodology is appropriate and complete but lacking an identification of limitations OR methodology is well designed but implemented poorly
• Very good – methodology is well suited, well implemented and a full discussion given of its limitations
• Excellent – a model for this standard that could be used for sample methodology for the evaluations sectoral/topical area

12. A complete description for stakeholder participation in the evaluation was given.

WHAT is the standard

The degree of participation of stakeholders in the evaluation process can vary along a continuum from low to high based on what key steps or activities stakeholders are involved in – some steps are more pivotal than others in shaping results – and what role stakeholders can have in each. Roles might include liaison, technical advisory roles, observer roles, or more active decision-making roles. The degree of participation of stakeholders should be described including why that particular degree was selected.

While not all evaluations can be participatory to the same degree, it is important that consideration is given to participation of stakeholders and that the evaluation report is transparent about the rationale and level of participation of different stakeholders.

WHY is this standard included

The participation of a broader range of stakeholders in an evaluation is increasingly recognised as a critical factor in the use of conclusions, recommendations and lessons. A human rights-based approach to programming adds emphasis to the participation of primary stakeholders. For UNICEF this clearly points to the involvement of children and adolescents. Finally, including certain groups of stakeholders may be necessary for a complete and fair assessment.

HOW is this standard applied
• Missing – no description of stakeholder participation is given (although stakeholders may be involved in the evaluation as data sources)
• Poor – vague description of stakeholder participation is given; OR a very limited definition of “participation” or “stakeholder” is used, e.g. no attention to primary stakeholders or participation is interpreted as involving stakeholders only as a source for data
• Satisfactory – clear description of stakeholder participation but no rationale provided for the degree of participation decided upon
• Very good – clear description and justification of stakeholder participation given, including reference to primary stakeholders
• Excellent – a model for this standard that exemplifies “participatory evaluation” with stakeholders included in evaluation design as well as data collection and recommendations workshop

LIMITATIONS to the use of this standard

The breadth and degree of stakeholder participation feasible in evaluation activities will depend in part on the kind of participation achieved in the programme/project. Nonetheless, evaluation activities can be used to open greater participation. Ideally there will be a few strategically important evaluation activities where a broader range of stakeholders can be brought together to explore common research/evaluation questions.

Given how dependent such participation is to the particulars of each programme/project, the reader applying this standard should generally look for soundness of the description and rationale given for the degree of stakeholder participation in the evaluation as opposed to rating the degree of participation itself. The Notes section of the rating sheet can be used to identify cases where a higher degree of participation may have been feasible and preferable – however this should not influence the rank on this standard.

***13. Where information was gathered from those who benefited from the programme/project, information was also gathered from eligible persons not reached.***

***If the methodology was designed such that direct beneficiaries were not contacted, this standard should be skipped. The previous standard covers the case that direct beneficiaries could have been included but were not.***

WHAT is the standard

To provide a complete assessment, people who were eligible to participate in or benefit from the programme/project but did not should be included in the evaluation. This includes children or adolescents who were eligible to benefit from the programme/project.

The decision to gather this information should be made with consideration of the ethical issues involved as outlined in the next standard. Although it may be technically feasible to collect information, it may not be ethically prudent.

WHY is this standard included

This standard is to assess how and why the programme/project did not attain its full potential. It is not to be confused with a “control group” – evaluation methodology that compares the programme/project area to an area explicitly not involved in the programme/project by design.

HOW is this standard applied
• Missing – information was gathered from primary stakeholders, but no effort was made to contact eligible persons not reached by the programme/project
• Poor – limited information was gathered from potential beneficiaries that does not allow for a full analysis of the programme/project and no explanation was given as to why it was so limited
• Satisfactory – adequate and appropriate data was collected from potential beneficiaries with some analysis given of the differences OR an adequate rationale given for why the information could not be gathered
• Very good – complete assessment of how and perhaps even why the programme/project did not reach the entire target population OR an adequate rationale given for why the information could not be gathered and suggestions for how to address this issue in the next programme/project cycle
• Excellent – a model for this standard that will assist future programme/projects to reach the entire target population

LIMITATIONS to the use of this standard

The context of the evaluation may limit the evaluator’s ability to acquire this information. If a sound rationale is given for why this analysis was not included, the report should not be marked down on this standard. This is not to deny the importance of this information but rather acknowledge the realities of some difficult situations such as unstable areas.

14. The evaluation design was ethical and included ethical safeguards where appropriate, including protection of the confidentiality, dignity, rights and welfare of human subjects, particularly children, and respect for the values of the beneficiary communities.

WHAT is the standard

Poorly designed efforts to expand the participation of primary stakeholders can do more harm than good. Working to expand the participation of any group of people engenders responsibilities to protect them from coercion to participate, from the negative repercussions of their participation, such as retribution by other stakeholders, and from other forms of abuse and exploitation. When the topic of the evaluation itself is contentious, there is a heightened need to protect those involved in the evaluation.

An evaluation report should contain a description of the measures and mechanisms put in place to: ensure that the evaluation process was ethical, that stakeholders were protected and address any ethical dilemmas or issues that emerged. This is especially important when children and adolescents are involved. Technical Note #1 "Children Participating in Research, Monitoring And Evaluation," available on the UNICEF Intranet discusses how to include children in an evaluation while respecting their rights and dignity.

Measures and mechanisms would include some process of ethical review of the design initially. Further the design should contemplate:
− the balance of cost and benefits to participants including potential negative impact,
− the ethics of who is included and excluded in the evaluation and how this is done,
− handling of privacy and confidentiality,
− practices of obtaining informed consent,
− feedback to participants and

16 http://www.intranet.unicef.org/epp/evalsite.nsf/1565f9b3780158a285256b95005a5231/acf4c8b740fa19c085256bad007a9bd9?OpenDocument
mechanisms for shaping and monitoring the behaviour and practice of evaluators and data collectors.

WHY is this standard included

Individuals have personal rights that are secured by law, by ethical practices, and by common sense and decency. Evaluators and evaluation managers have the responsibility to ensure that the rights and well-being of individuals not be affected negatively in planning and carrying out an evaluation. Lack of attention to protecting peoples’ rights and dignity is not only unethical but most often leads to unfair, inaccurate and/or incomplete evaluation results.

HOW is this standard applied

- Missing
- Poor – scant attention and only token paragraph on ethical considerations
- Satisfactory – good description of ethical considerations
- Very good – good description of ethical considerations including the rationale behind the design and the mechanisms to protect participants where appropriate
- Excellent – a model for this standard that could be referenced in guidance on this issue

15. In presenting the findings, inputs, outputs, and, where possible, outcomes/impacts were measured (or an appropriate rationale given why not).

WHAT is the standard

Findings regarding inputs for the completion of activities or process achievements should be distinguished clearly from results, and findings on results should clearly distinguish outputs, outcomes and where possible impact.

Outputs, outcomes and impacts should include any unintended effects – whether beneficial or harmful. Additionally, any multiplier or downstream effects of the programme/project should be included.

To the extent possible, each of these should be measured – either quantitatively or qualitatively. In using such measurements, benchmarks should be referred to as described in methodology.

WHY is this standard included

Using a framework distinguishing inputs, outputs and outcomes clarifies the various findings for the reader. Such clear classification is an essential element of results based management. For conclusions regarding attribution of results - which are most often tentative and require clear detailing of what is known and what can plausibly be assumed - it makes the logic from findings to conclusions more transparent and increases their credibility.

HOW is this standard applied

- Missing - blurring of activities, inputs, outputs, outcomes; no distinction made between them and no sense of the magnitude of the results
- Poor – some effort at measurement but a confusion between implementation and results, i.e. considers an activity an outcome
- Satisfactory - some logical distinction made in the findings showing the progression from implementation to results (though perhaps not labelled in the exact terms) but: with weak measurement OR with inadequate analysis of links in the results chain OR no rational given as to why no results analysis was given
• Very good - logical distinction made in the findings showing the progression from implementation to results (though perhaps not labelled in the exact terms) with appropriate measurement and analysis of the results chain or a rational given as to why an analysis of results was not provided
• Excellent – a model for this standard that shows a clear picture from the inputs/activities provided by various stakeholders to the outputs and, where possible, outcomes and results

LIMITATIONS to the use of this standard

Different terminology other than “input”, “output”, “outcomes”, or “impact” may be used. This is acceptable as long as a clear distinction is made between the different types of results and that activities are not viewed as a results. The language and classification used should illustrate a results based approach to analysis.

TIPS for report preparation

Data does not need to be presented in full; only data that supports a finding needs to be given, full data can be put in an Annex. Poor reports list data in one section and then presented findings in another section leaving the reader to juggle back and forth between the two sections to see if the findings are supported by the data. Additionally, reports should not segregate findings by data source. Excellent reports present a conclusion, then give findings and data to support that conclusion from all sources.

Findings should also cover all of the evaluation objectives. A report could have solid evaluation objectives and methodology but then the findings do not address all of the objectives using the data collected. This would mean the report is weak in Coverage for this section and could not be considered Excellent.

16. To the extent feasible, the report includes cost analysis.

WHAT is the standard

Using a range of cost analysis approaches – from the elaborate cost-effectiveness and cost-benefit analysis, to cost-efficiency analysis to the less orthodox ‘quick-and-dirty’ cost comparisons -- an evaluation can pursue the following broad questions:

- How do actual costs compare to other similar programmes or standard benchmarks?
- What is the cheapest or most efficient way to get the expected programme results?
- What will be the cost implications of scaling up or down?
- What will be the costs of replicating the programme in a different environment?
- Is this programme worth doing? Do economic benefits outweigh the costs?
- How do costs affect the sustainability of the programme?

Cost analysis in evaluation builds on financial information, but may also involve calculating “economic costs” such as human resources, labour-in-kind, opportunity costs, etc.

The scope of cost analysis, i.e. whether cost comparison is made concerning impacts, outcomes or outputs, will depend on the purpose of the evaluation and the evaluation questions posed. Cost analysis must be explicit in terms of the different perspectives from which cost are analysed (donors, a single organisation, primary stakeholders) and the limitations – the complexity of the programme (multiple programme objectives, partners, financial systems), the availability of data and the time and resources invested.
Cost analysis is not always feasible. It requires some specialized skills and the availability of appropriate data. Where no cost analysis is included in an evaluation, some rationale for this exclusion should be included in the objectives or methodology section.

It is also expected – though should not be confused with a full cost analysis – that evaluators point out areas of obvious inefficient use of resources.

**WHY is this standard included**

Addressing the evaluation criteria of efficiency requires some cost analysis. Additionally, satisfactory cost analysis strengthens results-based management and thus increases the utility of the evaluation. No cost analysis leaves significant questions unanswered.

**HOW is this standard applied**

- Missing – there is no cost analysis and no justification of the omission
- Poor – data is insufficient to provide analysis but conclusions are drawn from data and presented as the results of cost-analysis; OR data appears to be sufficient but there is no cost analysis and the justification for not addressing this issue is weak
- Satisfactory – cost analysis is provided that is well grounded in the findings; OR report provides reasonable and seemingly accurate justification for the exclusion with reference to availability of data and/or accepted limitations in the scope of the evaluation
- Very good – cost analysis is provided that is well grounded in the findings including an analysis of its limitations and recommendations for data collection to improve the situation for the next evaluation; OR report provides reasonable and seemingly accurate justification for the exclusion with reference to availability of data and/or accepted limitations in the scope of the evaluation and recommendations for data collection to improve the situation for the next evaluation are given
- Excellent – a model for this standard that could be used as sample methodology for other evaluations

**LIMITATIONS to the use of this standard**

There are numerous constraints to cost analysis. There may be a good rationale that is simply not given in the report. This again is a standard that all evaluations should strive to achieve and it is hoped that the organization will build capacity and data collection tools in the future.

**17. Analysis includes appropriate discussion of the relative contributions of stakeholders to results.**

**WHAT is the standard**

For results attributed to the programme/project, the result should be related back to the contributions of different stakeholders accurately. There should be a sense of proportionality between the relative contributions of each and the results observed.

**WHY is this standard included**

This is an integral element of results-based management and accountability to partners, donors and primary stakeholders.

**HOW is this standard applied**

- Missing – no discussion of stakeholders’ contributions in the findings/results section
• Poor – over-simplification of the relative contribution of stakeholders to the results OR report describes what each partner did (i.e. partners inputs/activities) but does not relate this to any output or outcome
• Satisfactory – a reasonable effort to distinguish the effect of the various contributions of stakeholders, though still fairly general and focusing on a few major stakeholders
• Very good – a fair and realistic effort to distinguish the correlation between the results and the relative contributions of stakeholders
• Excellent – a model for this standard that provides a comprehensive analysis

LIMITATIONS to the use of this standard

Although difficult and perhaps impossible given the methodological constraints of certain evaluations, this is a standard that all evaluations should strive to achieve. A rationale for why such analysis was not done should be provided in the report if it cannot be completed. If an evaluation is done early in the programme/project cycle, it is understandable that results or any link to a stakeholders contribution may not be found.

18. Reasons for accomplishments and difficulties of the programme/project, especially continuing constraints, were identified as much as possible.

WHAT is the standard

Evaluation go beyond a description of programme/project implementation and outcomes to “why”. An analysis of the underlying causes, constraints and opportunities should be given. External factors contributing to the accomplishments and difficulties should be identified and analyzed to the extent possible. Beyond simply describing the geographic and demographic characteristics, the social, political, or environmental situation that has affected the outcome of the programme/project should be assessed. Informed judgments about what results may reasonably be attributed to the intervention, and what results may be due to other factors should be provided.

WHY is this standard included

Without an analysis of the reasons, an evaluation may wrongly attribute success/failure to something which is not related to the programme/project leading to inaccurate findings and recommendations. An explanation of context contributes to the utility and accuracy of the evaluation. Additionally, an understanding of which external factors contributed to the success or failure of a programme/project allows the programme/projects to be replicated elsewhere.

HOW is this standard applied

• Missing – findings are presented without any reasons
• Poor – reasons are identified but seem unrelated to the findings
• Satisfactory – reasons are identified and seem to logically flow from the findings
• Very good – reasons are identified that are based on the findings and analysis of how such factors will affect the future of the programme/project (or generally for other programmes/projects) are given
• Excellent – a model for this standard providing concise, usable information, easily accessible in the report that can improve programming in the next phase

LIMITATIONS to the use of this standard
Since the reader is removed from the context him/herself, the reader cannot access the accuracy of the evaluator's description and instead should focus on the completeness and logical consistency of the description provided.

Sometimes the data that can be gathered does not provide this information. However it is extremely helpful in designing better projects/programs in the future. A certain amount of reasonable speculation on the part of the evaluator is welcome if clearly identified as conjecture and well grounded in the findings presented.

This standard refers to findings – investigating why. The standard below is somewhat related – it refers to the conclusions drawn from these findings and that they should be insightful.

19. **Conclusions were substantiated by findings consistent with data and methods and represent insights into identification and/or solutions of important problems or issues.**

**WHAT is the standard**

Conclusions should add value to the findings. Users of evaluations must clearly see the logic of conclusions and how this flows from the actual findings. With this, they are then able to accept or reject additional analysis and reflections of the evaluators.

Conclusions must also focus on issues of significance to a programme/project. This choice of significant issues must relate back to the evaluation objective and key questions the evaluation is trying to answer. Simple conclusions that are already well known and obvious are not useful and should be avoided.

**WHY is this standard included**

Users must be able to discern the logic of the conclusions, be convinced of the conclusions before they will accept them.

**HOW is this standard applied**

- Missing – raw data presented with findings simply restating the data totals, lacking conclusions
- Poor – explanation of results not related to actual data; OR conclusions given that do not follow from the findings presented
- Satisfactory – explanation of results consistent with data and conclusions well based on findings
- Very good – explanation of results consistent with data, conclusions well based on findings, conclusions represent actual insights into identification and/or solutions of important problems or issues
- Excellent – a model for this standard providing concise, usable information, easily accessible in the report that can improve programming in the next phase

**LIMITATIONS to the use of this standard**

When conclusions do not correlate to the findings, the reader does not know what exactly the conclusion was based on - a case of insufficient description by the writer or perhaps a point of view that has found its way into the evaluation as a fact. The conclusions section may get a poor score because of a poor description, despite the quality of the analysis.

20. **Recommendations were firmly based on evidence and analysis; they were directly relevant and realistic with priorities for action made clear.**
WHAT is the standard

Recommendations are the logical implications of the findings and conclusions. They should be relevant to the programme/project, not broad or vague, and realistic.

At the same time, recommendations that are overly specific or represent a long list of items are also of little value. Changes to the programme/project should be part of a participatory process that involves UNICEF staff considering the recommendations through a planning process – not the sole determination of someone outside the organization.

The preparation of recommendations needs to suit the evaluation process. Participation by stakeholders in the development of recommendations is strongly encouraged to increase ownership and utility. The planners and managers of the evaluation may decide to: include stakeholders in the creation of recommendations presented in the report; or may leave the consultation process for a separate stage after the report is completed - meaning that recommendations included in the evaluation report will necessarily be less specific, perhaps called “implications” rather than “recommendations”. The description of the evaluation methodology and discussion of participation would alert the reader to this situation. In such cases, the implications can still be assessed as to how logically they flow from findings and conclusions.

WHY is this standard included

For accuracy and credibility, recommendations should be related to the findings and conclusions.

HOW is this standard applied

- Missing
- Poor – recommendations or implications are given that do not flow logically from the findings; OR the consultant has given implications from the findings that are well-grounded but no follow-up process is described
- Satisfactory – recommendations are well-grounded in the findings, but are either: not specific and relevant, not few in number (e.g. 10 – 15) or not explicitly prioritised; OR the consultant has given implications from the findings that are well-grounded and explains the follow-up process planned with UNICEF staff
- Very good – recommendations are well-grounded in the findings and are either few in number (e.g. 10 – 15) or explicitly prioritised OR the consultant has given implications from the findings that are well-grounded and UNICEF staff have considered the implications and provided some type of follow-up documentation
- Excellent – a model for this standard, the evaluation or follow-up documentation provide clear, relevant, credible, prioritized, insightful information

TIPS for report preparation

Sometimes evaluators present recommendations within their conclusions. As long as the recommendations are clearly identified as recommendations and some how distinguished, e.g. in italics or a box, this is acceptable. Mixing recommendations with conclusions leads to confusion – it is hard to know if the evaluator is truly recommending action or simply a poor choice of wording; also recommendations are not prioritized and the reader has an unclear picture of what the organization’s next steps should be. This would result in a “Poor” rating.

***21. Lessons learned, when presented, were generalized beyond the immediate intervention being evaluated to indicate what wider relevance there might be.***
***Not all evaluations have lessons learned. If a report does not identify any lessons learned, this standard should be skipped. For a reader removed from the evaluation process, conclusions that could be lessons learned that are not identified as such cannot be identified with any methodological rigor. (If a reader can determine a lesson learned, it is most likely a commonly accepted lesson that would not contribute any new knowledge.)***

**WHAT is the standard**

Lessons learned are contributions to general knowledge. They should be well supported by the findings and conclusions presented. They may refine or add to commonly accepted lessons learned but should not be merely a repetition of common knowledge. Recommendations are different from lessons learned in that they are specific and relevant to the particular programme/project in its specific context.

**WHY is this standard included**

This increases the usability of the report contributing to learning outside the programme/project, at the country level, regionally or globally, within the organisations involved or beyond.

**HOW is this standard applied**

- Poor – findings or conclusions are inaccurately identified as lessons learned OR lessons learned are identified that are simply repetitions of basic knowledge about good programming
- Satisfactory – lessons learned are correctly identified and stem logically from the findings
- Very good – lessons learned are correctly identified, stem logically from the findings, and an analysis of how they can be applied to different context and/or different sectors is given
- Excellent – a model for this standard; well written and described such that the lessons could be simply lifted from the report and put into a newsletter/update

Note: a description of MISSING is not provided since this standard is not to be applied if there are no lessons learned identified. See the note at the beginning of this standard.

**LIMITATIONS to the use of this standard**

Some reports identify lessons learned that are truly just findings that can related to more than one programme/project within that context – but not necessarily to a different context or a different sector. In this case, the rating would be Poor.

**22. Completeness of Annexes**

**WHAT is the standard**

Additional supplemental information to the evaluation that should be included in the Annexes includes:

- List of persons interviewed and sites visited. (Maps, especially DevInfo maps are helpful but not required.)
- Data collection instruments (copies of questionnaires, surveys, etc.)
- Terms of reference.
These Standards do not analyze the quality of the TOR. Additional information on TORs can be found in Technical Note 2 “What goes into a Terms of Reference” on the Intranet.

- List of Abbreviations.
- Cost of evaluation is given, preferable presented as a percentage of the overall programme/project cost. Costs should include an accounting of the use of staff time and other UNICEF CO resources.
  Although not an audit or meant for audit purposes, this information is provided to promote transparency and ensure propriety. This can also assist the CO in tracking evaluation costs and in planning the next evaluation.

**WHY is this standard included**

The annexes increase the usability of the report – other COs often look at others’ survey questions and data collection instruments when designing their evaluations.

**HOW is this standard applied**

A simple checklist is used to see if the information was included or not.

**LIMITATIONS to the use of this standard**

When these Standards are used at the headquarters level, the full annexes may not have been sent with the report because of space limitations on the size of electronic attachments. The reader is asked to check the Table of Contents to see if the Annexes are described there. If the Table of Contents lists the contents of the annexes, the reader will mark the various contents of the annexes listed as included.

**TIPS for report preparation**

Completing a checklist before final payment is given to the consultant will ensure that all the necessary material is present.

Administrative note for the COs - If the Annexes are too long to include with the report, COs are asked to send the Annex to Lourdes SanAgustin via email. Please tell her the name of the report and its sequence number. Our goal is to have the TOR for every evaluation that is included in the ERD on the Intranet for staff to use as a reference.

---

17 http://www.intranet.unicef.org/epp/evalsite.nsf/1565f9b3780158a285256b95005a5231/e17ce448d105c58485256b88006e824c?OpenDocument
III. Glossary of Terms

**Activity(ies)**
Actions taken or work performed through which inputs - such as funds, technical assistance and other types of resources – are mobilized to produce specific outputs. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management, 2002)

**Coherence**
Refers to the policy coherence and the need to assess security, developmental, trade and military policies to ensure that there is consistency and, in particular, that all policies take into account humanitarian and human rights considerations. (ALNAP Annual Review Glossary, 2003)

**Conclusions**
Conclusions point out the factors of success and failure of the evaluated intervention, with special attention paid to the intended and unintended results and impacts, and more generally to any other strength or weakness. A conclusion draws on data collection and analyses undertaken through a transparent chain of arguments. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management, 2002)

**Context (of an evaluation)**
The combination of factors accompanying the study that may have influenced its results, including geographic location, timing, political and social climate, economic conditions, and other relevant professional activities in progress at the same time. (Programme Policy and Procedures Manual, UNICEF, May 2003)

**Cost Benefit Analysis**
Cost-benefit looks to outcomes and impact, but compares different interventions with different effects. It compares the benefits-to-costs ratio, which is the total monetary value of benefits compared to total monetary value of costs.

Because cost benefit analysis involves translating all inputs and all outcomes into a common unit of comparison (e.g. dollars), it is then possible to compare programmes with different objectives. For example, two programmes with the goal of poverty alleviation, but different objectives — e.g. expanded access to credit for investment in cash crops vs. access to vocational training to meet needs in local processing industries — can be compared. (M&E Training Resources, UNICEF, 2004)

**Cost Effectiveness Analysis**
Cost-effectiveness analysis entails comparing costs across different strategies for achieving a given outcome, with a view to determining the lowest cost approach. For example, cost-effectiveness analysis might explore three different approaches to getting girls working in the informal sector back into school. As compared to cost-efficiency analysis, it is wider in scope, looking beyond outputs to outcomes. (M&E Training Resources, UNICEF, 2004)

**Cost Efficiency Analysis**
Cost-efficiency analysis compares costs of how project inputs are supplied and used to achieve specific outputs with a view to finding lowest cost options. For example, cost-efficiency analysis might explore whether outreach vaccination services were provided at the lowest possible cost. It entails comparing total and unit costs (total cost/number of units of outputs) as well as comparing
cost breakdowns among different sites or facilities, or over time at the same site. It entails exploring what makes costs higher and why, and takes into consideration quality of service.

Cost-efficiency analysis tends to have a smaller scope of analysis, comparing cost at the level of outputs. It does not relate costs to broader issues of outcomes as does cost-effectiveness analysis. Efficiency is a necessary and not a sufficient condition for effectiveness. (M&E Training Resources, UNICEF, 2004)

Coverage
The need to reach major population groups facing life-threatening suffering wherever they are, providing them with assistance and protection proportionate to their need and devoid of extraneous political agenda. (ALNAP Annual Review Glossary, 2003)

Effectiveness

Efficiency
An economic term referring to the measure of the relative cost of resources used in a programme to achieve its objectives. (Programme Policy and Procedures Manual, UNICEF, May 2003)

Impact
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. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management Proposed Harmonized Terminology, 2002)

Input
The financial, human, material, technological and information resources used for the development intervention. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management Proposed Harmonized Terminology, 2002)

Lesson learned
Conclusions that can be generalized beyond the specific case. This could include lessons that are of relevance more broadly within the country situation or globally, to an organization or the broader international community. (Programme Policy and Procedures Manual, UNICEF, May 2003)

Outcome
The intended or achieved short-term 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. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management Proposed Harmonized Terminology, 2002)
The products and services which result from the completion of activities within a development intervention. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management Proposed Harmonized Terminology, 2002)

Programme
In UNICEF, "programme" is used in two ways: 1. "country support programme" defined as the whole field in which UNICEF is co-operating in the country, e.g. the subject of a programme recommendation to the Board; and 2. co-operation in a sector or a geographical area, e.g. a "health programme" or an "area-based programme." Programmes are designed to have a specified outcome(s) or impact, and are detailed in a Plan of Operations, Master Plan of Operations (MPO) or Country Programme Action Plan (CPAP), if it consists of a number of programmes.

Frequently a programme consists of a set of projects, which in turn are made up of activities. A project is usually related to one main implementing agency, therefore to one sector (e.g. health) or part of a sector (e.g. immunization) or one field (e.g. women's activities). It can be defined as "a planned undertaking composed of a group of interrelated activities whereby specified inputs are designed with the purpose of achieving specified outputs (or changes) within a specified time-frame." (Programme Policy and Procedures Manual, UNICEF, May 2003)

Project
A time-bound intervention that consists of a set of planned, interrelated activities aimed at achieving defined objectives. (Programme Policy and Procedures Manual, UNICEF, May 2003)

Propriety
The evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results. (Program Evaluation Standards, American National Standards Institute, 1989)

Protection
Activities aimed at obtaining full respect for the rights of the individual in accordance with the letter and the spirit of the relevant bodies of law (i.e. human rights, humanitarian and refugee law) which are conducted impartially and not on the basis of race, national or ethnic origin, language or gender. (ALNAP Annual Review Glossary, 2003)

Recommendation
Prescription of what should be done, in the future and in a specific situation with regard to a programme, project, strategy or policy under study. (Programme Policy and Procedures Manual, UNICEF, May 2003)

Relevance
The extent to which the objectives of a development intervention are consistent with duty bearers and rights holders requirements, country needs, global priorities and partners’ and donors’ policies.

Retrospectively, the question of relevance often becomes a question as to whether the objectives of an intervention or its design are still appropriate given changed circumstances. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management, 2002)
**Result**
Results are changes in a state or condition which derive from a cause-and-effect relationship. There are three types of such changes (intended or unintended, position and/or negative) which can be set in motion by a development intervention – its output, outcome and impact. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management Proposed Harmonized Terminology, 2002)

**Results Based Management (RBM)**
A management strategy by which an organization ensures that its processes, products and services contribute to the achievement of desired results (outputs, outcomes and impacts). RBM rests on clearly defined accountability for results, and requires monitoring and self-assessment of progress towards results, and reporting on performance. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management Proposed Harmonized Terminology, 2002)

**Results chain**
The causal sequence for a development intervention that stipulates the necessary sequence to achieve desired objectives beginning with inputs, moving through activities and outputs, and culminating in outcomes, impacts and feedback. It is based on a theory of change, including underlying assumptions. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management Proposed Harmonized Terminology, 2002)

**Stakeholder**
Individuals, groups, or organizations that can affect or be affected by an intervention or issue. **Primary stakeholders** are those directly benefiting from an intervention. (Programme Policy and Procedures Manual, UNICEF, May 2003)

**Sustainability**
The continuation of benefits from a development intervention after major development assistance has been completed.

The probability of long-term benefits. The resilience to risk of the net benefit flows over time. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management, 2002)

**Utility**
The extent to which an evaluation produces and disseminates reports that inform relevant audiences and have beneficial impact on their work. One of the key Standards against which an evaluation is measured. (Programme Policy and Procedures Manual, UNICEF, May 2003)
Children Participating in Research, Monitoring And Evaluation (M&E) — Ethics and Your Responsibilities as a Manager

The trend of involving children more actively in M&E programmes as part of their right to participate brings many practical challenges and raises ethical considerations.

The ethical issues are complex and no straightforward guidelines exist. Children’s rights are established in international law, where children are defined as those up to 18 years old, but the reality and meaning of childhood throughout the world differs. The context (cultural, political etc.), the capacities of each child, which in turn vary with age and stage of development, and the corresponding possibilities for participation all vary; so, too, then, will the response to ethical challenges in research and M&E practice (Boyden and Ennew, 1997).

The responsibilities lie with researchers/evaluators, those technical professionals involved in design of research, monitoring and evaluation activities and directly in data collection. However, managers commissioning such activities are equally responsible for ensuring that ethical issues are identified and resolved in methodology design.

This Evaluation Technical Notearticle explores the child’s right to participate, related concepts and their implications for research and M&E activities. It also outlines key ethical considerations about whether and how children participate in research and M&E, and provides a checklist of questions for research and M&E managers.

Guidance from the Convention

The Convention on the Rights of the Child provides clear initial guidance for children’s participation in programmes, and in research and M&E:

- All rights guaranteed by the convention must be available to all children without discrimination of any kind. Equity and non-discrimination are emphasised.
- The best interests of the child must be a major factor in all actions concerning children. This puts the onus on researchers and evaluators who encourage children’s participation to consider carefully how this supports the best interest of each child.
- Children’s views must be considered and taken into account in all matters that affect them. They should not be used merely as data from subjects of investigation.

The four articles related to participation further establish the parameters:

- Article 12 states that children who can form their own views should have the right to express those views and have them taken into account. However, the right to participate and freedom of expression are not equated with self-determination. Each child’s views are their “reality”, which must be considered, but also must be weighed against the best interests of the child in any decisions eventually taken.
• **Article 13** states that children have the right to freedom of expression, which includes seeking, receiving and giving information and ideas through speaking, writing or in print, through art or any other media of the child's choice. Their participation is not a mere formality; children must be fully informed and must understand the consequences and impact of expressing their opinions. The corollary is that children are free to not participate, and should not be pressured. Participation is a right, not an obligation.

• **Article 14** establishes that State parties must respect children's right to freedom of thought, conscience and religion, as well as parents' or guardians' role in their exercising this right. Research and M&E activities seeking to involve children must clearly acknowledge and ideally seek to build on these respective roles.

• **Article 15** establishes that the States parties must recognise children's right to freedom of association and of peaceful assembly. As children's capacities evolve, they will increasingly participate and seek the representation of their perspectives in wider fora — at community, sub-national, national and global levels. Research and M&E activities can help this evolution along.

The Convention establishes that participation should be seen as both a process and an end in and of itself; that the very act of participation should be seen as contributing to the development of the children involved. This suggests highly participatory approaches to research and M&E where children are involved from design to the use of results.

**What is participation?**

While the Convention establishes a right to participate, M&E experience shows that "participation" is many things to many people — true for the participation of adults as much as for children. "Participatory" approaches to M&E range from those that survey the opinions of "beneficiaries" or primary stakeholders, to those where primary stakeholders are placed at the centre of the process, from design to implementation, to analysis and follow-up of M&E exercises. The premises and limitations of the model and methodology must be clearly stated — we must call it what it is.

Several different angles can be taken to define the nature of children's participation. Roger Hart (Hart, 1992) used an eight-degree scale:

<table>
<thead>
<tr>
<th>Degrees of participation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Child-initiated, shared decisions with adults</td>
<td></td>
</tr>
<tr>
<td>- Child initiated and child-directed projects</td>
<td></td>
</tr>
<tr>
<td>- Adult-initiated, sharing decisions with children</td>
<td></td>
</tr>
<tr>
<td>- Participation in which children are consulted and informed (run by adults, but children understand the process and their opinions are treated seriously).</td>
<td></td>
</tr>
<tr>
<td>- Assigned but informed participation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Non Participation</td>
<td></td>
</tr>
<tr>
<td>- Tokenism...children are given a voice but have little choice about the subject, the style of communicating it or any say in organising the occasion</td>
<td></td>
</tr>
<tr>
<td>- Decoration ... children are asked to take part in an event but are not given any explanation of the issues or the reason for their involvement</td>
<td></td>
</tr>
<tr>
<td>- Manipulation</td>
<td></td>
</tr>
</tbody>
</table>
Efforts that fall under tokenism, decoration and manipulation not only fail in their objective to foster the participation of children, but can also discredit the effort and the organisations involved, ultimately undermining the meaning of the right to participate.

This ladder includes the relationship between children and adults (Rajani, 2000), be they programme managers or researchers, which is important. The increasing degrees of participation suggest increasingly evolved capacities of children and corresponding capabilities of adults towards encouraging the participation of children.

Context is also important. Political, social and economic contexts will have their own institutional norms and practices at different levels (national, sub-national, community, family), and in different fora will favour (or limit) participation to different degrees. Analysing context can reveal how it limits participation, as well as how participation can be increased.

Rakesh Rajani’s “Framework for Promoting Effective Adolescent Participation” (see page 4) links the above two aspects — context and the relationship between children and adults — with other factors to define the nature of participation.

It illustrates children’s roles from listening to active decision-making roles, and the different spheres in which they participate, both in terms of geographical and institutional settings. Three key contributing factors underlie these facets of participation: the individual capabilities of children, the supporting environment and the opportunities created for participation. Programme interventions that strive to build children’s (adolescent’s) participation must do so by trying to influence and change these contributing factors.

“A Framework for Promoting Effective Adolescent Participation” (Rajani, 2000: 13)
These two frameworks are not only good for designing programmes, but for defining the participatory activities for research and M&E exercises as well, i.e. where children will participate, in what role and through what type of interaction with adults. If the M&E activity itself is designed to build participation, then managers and evaluators must specify how the activity will influence children’s capabilities and their supporting environment and therefore their opportunities for participation.

**Ethical issues**

Several complex ethical issues emerge around children’s participation in research and M&E without a guideline on how to respond to any of them. They include:

- **Accountability.** Since researchers and programme managers are accountable to a wide range of stakeholders (including primary stakeholders, i.e. those intended to benefit from programme interventions), and the involvement of primary stakeholders in research and M&E activities is an expression of this accountability, then research and M&E should also involve the participation of children. Their participation is relevant not only where planned interventions and issues specifically affect them, but also where they, as members of the wider community, are affected (e.g. in relation to safe drinking water). It must be clear in initial research and M&E design proposals what role will children play and how will they be involved.

- **Protection of children’s best interests.** This has very clear and powerful implications for the process of research and M&E as well as for the dissemination of its results.
  
  - Children must not be exposed to risks if there is no benefit to them. These include the psychological effects on the individual child of participating in the activity (for example, in cases of abuse where the fears and pain of past experiences re-emerge); the social costs of participating such as negative effects on family and community relations; more acute threats such as reprisals by people who feel threatened by children’s participation; and misuse of information, ranging from sensationalist media attention or to more sinister uses in situations of conflict and humanitarian crisis. Weighing these risks against possible benefits requires careful judgement, particularly where risks to individual children are done in the name of broad sometimes incremental societal changes.

  - Those leading and carrying out research and M&E activities are also responsible for protecting children from placing themselves at risk, even where a child might be willing to participate and voice their views (Boyd en and Ennew, 1997; Boyden 2000).

  - The responsibility to protect children may also entail withholding information from children where that information may place them at risk (Boyd en and Ennew, 1997). Children may not always be able to cope with the implications of information received or may not be able to judge adequately when and with whom to share that information.

  - Research and M&E activities must be able to ensure confidentiality. However, information may at times reveal that a child is at risk or is a risk to others, which is why design of research and M&E activities must include guidelines for breaking confidentiality and intervention, including defining what follow-up and referrals can be made. Children must be made aware of the limits to confidentiality and possible intervention based on what is in their best interests.
• **Informing children.** Research and M&E managers are responsible for ensuring that children receive the information they need to form and express their views as well as to decide whether they choose to express them at all. To "inform" should be understood as meaning more than simply providing information. How information is conveyed must be appropriate to the context and to children's capabilities. How informed children are affects how their views can be interpreted.

• **Informed consent.** The focus of most ethical guidelines is on research in the West, this has often involved signed consent forms to ensure that participants in research are aware of any potential implications of their involvement (by the same token to protect researchers from liability). Researchers must respect the consent regulations of the countries in which they are working, however, parental consent is not an adequate standard in light of the rights of the child. Informing children of the potential implications is required. Further, consent should not be a one-time event in the course of a child's participation. It should be a negotiation of the parameters and limits of his/her participation, an ongoing exchange in which a child's views and best interests are paramount (Alderson, 1995). All issues of negotiating consent and encouraging children to express themselves must be carried out with clear recognition of the natural power imbalance between children and adults.

• **Equity and non-discrimination.** Those involved in research and M&E must ensure that selection of those children who participate and the processes and methods used serve to correct, not reinforce, patterns of exclusion. This requires attention to socio-economic barriers including gender and age discrimination as well as to the different ways and capacities in which children express themselves.

• **Respect of children and their views.** Those involved in design must choose methods and processes that best facilitate children expressing their views. Methods will most often be qualitative, and processes will likely be capacity building or participatory. However, respecting children's views does not mean allowing them to dictate conclusions. A child's input, like that of any other stakeholder, must be weighed as one perspective and interpreted in light of his/her experience, interests and evolving capacities. Assumptions and frameworks for interpreting information must be appropriate to the children involved and transparent to ensure credibility with users of research and M&E results.

• **Ownership.** Children must be informed of the results of the research. And since children will likely express themselves by diagramming and drawing, they should also be given rights of ownership of the research "data".

• **Methodological limitations.** It is unethical to carry out data collection if the design will not achieve the research objectives or respond to evaluation questions. Methodological limitations must be considered carefully, including the potential effects of power relations between children and adults. In order to increase children's participation, methodologies will likely tend towards the more qualitative with more specific adaptations for the children involved, and findings will be representative of narrower populations. Those involved in initial research and M&E design must balance degrees of participation of children with the credibility and breadth of application of research and M&E results.
Questions For Managers Of Monitoring, Evaluation & Research Activities:

The following is adapted\(^1\) from P. Alderson (1995), "Listening to Children: Children, ethics and social research", Barnardos, primarily from "Ten Topics in Ethical Research" (p.2-6) with detailed extracts on key issues in boxed text. While the original questions refer to research, they are equally relevant for UNICEF monitoring and evaluation work.

UNICEF offices are responsible for ensuring that these questions are considered in the design of the monitoring, evaluation and research activities in which they are involved.

1. Purpose

- Is the topic worthwhile? How are the findings likely to benefit children? How will they add to what is already known?
- If the findings are meant to benefit certain children, who are they and how might they benefit?
- Assuming findings are to be used to facilitate decision-making, who do they target? Is children's role in decision-making facilitated by this activity?

2. Costs and hoped-for benefits

- What contributions are children asked to make, such as activities or responses to be tested, observed or recorded? Is this a one-off contribution or, as in the case of some monitoring activities, will this be repeated?
- Might there be risks or costs — time, inconvenience, embarrassment, intrusion of privacy, sense of failure or coercion, fear of admitting anxiety? Also, consider retribution in contexts of conflict.
- Might there be benefits for children who take part — satisfaction, increased confidence or knowledge, time to talk to an attentive listener, an increased role in decision-making processes affecting them?
- Are there risks and costs if the research, monitoring or evaluation activity is not carried out?
- How can the researchers or managers of research and M&E promote possible benefits of the work?

"Are attempts made to avoid or reduce harms? Such as rehearsing with children a way of saying 'no' when they do not want to reply, assuring them that this will be respected and they will not be questioned about why they say 'no', or ensuring that children who feel worried or upset about the research can talk to someone about it afterwards? It can be useful to try to find out gently why young people want to refuse. Does the research seem boring or irrelevant? Could it be improved with their help?" (Alderson, 1995 -19)

---

\(^1\) Questions were rephrased and adapted, and a very few additions made, to apply to both monitoring and evaluation as well as to make the list more appropriate to developing country contexts. Some sections considered less relevant to UNICEF work have been deleted.
• What is planned to prevent or reduce any risks? What is the guidance regarding data collectors/researchers response to children who wish to refuse or withdraw? What will be the procedure with children who become distressed (e.g. if they simply feel uncomfortable, or if participation requires them to relive or experience emotional or psychological trauma) on the spot and in terms of referrals and follow-up? What steps are taken to ensure the protection and supervision of the children involved, including against bad practices by data collectors/researchers?

• Are the methods being tested with a pilot group? Will risks and costs be reassessed after piloting and what protection is offered to children involved in the pilot?

3. Privacy and confidentiality

• How will the names of children be obtained, and will they be told about the source?

• Does the selection method allow children and parents to opt into the activity (e.g. to volunteer for selection)? Is the selection method intrusive or coercive?

• Will interviews directly with individuals be conducted in a quiet, private place?

• Can parents be present or absent as the child prefers?

• In rare cases, if front line researchers/evaluators think that they must report a child’s confidences, such as when they think someone is in danger, will they try to discuss this first with the child? Do they warn all children that this might happen? Who will they report to and who/how many people will be involved? Who will guide this process?

• Will personal names be changed in records and in reports to hide the child’s identity? What should be done if children prefer to be named in reports?

• Will the data collection records, notes, tapes, films or videos, be kept in lockable storage space? Who will have access to these records, and be able to identify the children?

• When significant extracts from interviews are quoted in reports, should researchers/evaluators first check the quotation and commentary with the child or parent concerned? What should be done if respondents want the reports to be altered?

• Is there some verification that the field researchers in direct contact with the children do not represent a risk to children, i.e. have the appropriate values, attitudes and skills to deal with each child ethically and compassionately?

• Should records be destroyed when the research or M&E activity is completed or when related programme activity ends?

• Will the children be re-contacted at different points during the course of the programme for ongoing monitoring or evaluation, or is it ethical to ask the same children to take part in another research activity? In either case, how will the list of contact names be managed, stored?
4. **Selection, inclusion and exclusion**

- Why have the children concerned been selected to take part in the activity?
- Have efforts been made to reach marginalised, indigenous or disadvantaged children? Are issues of accessing these children satisfactorily dealt with in the methodology?
- If some of the children selected do belong to disadvantaged groups, have the researchers made allowance for any extra problems or anxieties they may have? Does the methodology accommodate their differing capacities?
- Have some children been excluded because, for example, they have speech or learning difficulties? Can the exclusion be justified?
- Are the findings intended to be representative or typical of a certain group of children? If so, have the children in the study been sufficiently well selected to support these claims?
- Do the design and planned numbers of children to be involved allow for refusals and withdrawals? If too many drop out, the effort may be wasted and therefore unethical. Consider also the possibility of withdrawals at different points in repeated monitoring activities.
- If the issue or questions being investigated are about children, is it acceptable only to include adult subjects?

5. **Funding**

- Are the children's and parents' or carers' expenses repaid?
- Should children be paid or given some reward after helping with the activity? Does the role of the children play a factor in whether or not they are paid, i.e. if children are active decision-makers as opposed to interviewees?
- How do these practices compare to those of other organisations working in the same region?
- How do the practices of paying children compare with payment of adults involved (e.g. parents, teachers, other community members)?

6. **Process of review and revision of ToRs and methodological proposal**

- Have children or their carers helped to plan or comment on the methodological proposal?
- Has a committee, a small group or an individual reviewed the protocol specifically for its ethical aspects and approach to children?
- Is the methodological design in any way unhelpful or unkind to children?
- Is there scope for taking account of comments and improving the design?
- Are the researchers accountable to anyone, to justify their work? Are researchers', managers' and other stakeholders' responsibilities vis-à-vis ethical practices clearly established?
- What are the agreed methods of dealing with complaints?
7. Informing children, parents and other carers

- Are the children and adults concerned given details about the purpose and nature of the research or M&E activity, the methods and timing, and the possible benefits, harms and outcomes? If children are not informed, how is this justified?
- Does a researcher/evaluator also encourage children and adults concerned to ask questions, working with an interpreter if necessary?
- If the research is about testing two or more services or products, are these explained as clearly and fully as possible?
- Are key concepts, such as 'consent', explained clearly?
- Are children and/or adults given a clearly written sheet or leaflet to keep, in their first language? If literacy is an issue, how is this handled in terms of ensuring children and their carers can access and review information provided about the activity at a later time?
- Does the leaflet give the names and address of the research/data collection/evaluation team? How can children contact someone from the team if they wish to comment, question or complain?

8. Consent

- As soon as they are old enough to understand, are children told that they can consent or refuse to take part in the activity?
- Do they know that they can ask questions, perhaps talk to other people, and ask for time before they decide whether to consent?
- Do they know that if they refuse or withdraw from the activity this will not be held against them in anyway?
- How do the researchers/evaluators help the children to know these things, and not to feel under pressure to give consent?
- How do they respect children who are too shy or upset to express their views freely?
- Are parents or guardians asked to give consent?
- How will the situation be handled if a child wants to volunteer but the parents refuse?
- Is the consent written, oral or implied? What is legally required and appropriate in the context?
- If children are not asked for their consent, how is this justified?
9. Dissemination

- Will the children and adults involved receive short reports on the main findings or other forms of feedback?
- Are the capacities of children and their preferences for how they receive feedback taken into consideration?

10. Impact on children

- Does the research, monitoring or evaluation activity have any impact on children’s capabilities, on the degree to which their environment is supportive of their participation (e.g. a change to attitudes of parents or other adults, to customs or to laws) or on future opportunities for participation (e.g. a change to practices in schools or other fora where children may participate; the creation of new fora, organisations etc.)? Was any such impact planned for in the design?
- Have children involved been realistically prepared for the expected impact, whether small or large?
- Besides the effects of the activity on the children involved, how might the conclusions affect larger groups of children?
- What models of childhood are assumed, e.g. children as weak, vulnerable and dependent on adults; as immature, irrational and unreliable; as capable of being mature moral agents; as consumers? How do these models affect the methods of collecting and analysing data.
- Is the approach reflexive, in that those involved in data collection and analysis critically discuss their own prejudices?
- Do they use positive images in reports and avoid stigmatising, discriminatory terms?
- Do they try to listen to children and in children’s own terms, while aware that children can only speak in public through channels designed by adults?
- Do they try to balance impartial assessment with respect for children’s worth and dignity?

"What will the intended and possible impact be on children? How will the research be done? And, in some cases, should it be done at all? These questions entail taking account of the status of children in society. An 'impact on children' statement for each research proposal would examine the likely effects of the research questions, methods and conclusions on the child subjects and on all young people affected by the findings. Will the research reinforce prejudice about children’s inabilities and faults by portraying them as victims or villains? Or will researchers examine these beliefs and devise methods which investigate children’s capacities and their needs and interests from the children's points of view? "

(Alderson, 1995 - 41)
References and further readings:

General


Design and methodology


What goes into a Terms of Reference (ToR)

A Terms of Reference (ToR) — also known as a Scope of Work — is a plan or blueprint outlining the key elements of the purpose, scope, process and products of an activity, including management and technical aspects as necessary.

Developing a ToR is a critical early step in any evaluation. In the narrowest sense, it is the basis for contractual arrangements with external consultants. More broadly, it should first be developed as a means of clarifying expectations, roles and responsibilities among different stakeholders, providing the plan for the overall activity, including follow-up. The time and effort spent in preparing a good ToR has big returns in terms of the quality, relevance and usefulness of the product.

The depth and details in the ToR will of course vary. The ToR for an externally facilitated programme evaluation involving numerous stakeholders will be quite detailed, while for an internal evaluation of an activity or an emergency rapid assessment it could be a simple outline.

ToRs are often developed in stages. In programme evaluation, stakeholders’ first discussions will focus on the details on purpose and evaluation questions. A further developed version used for recruiting external consultants requires more detail on existing information sources, team composition, procedures and products, but may describe methodology and a calendar of activities only in broad terms. The ToR may be further refined once an evaluation team is on board, with a careful review of the purpose and key questions and corresponding elaboration of methodology.

ToRs are important:

For all stakeholders
- They explain the agreed expectations in terms of the parameters and process of the exercise, and are a guide to each stakeholder’s specific role.

For the evaluation or assessment/survey team
- They ensure that expectations are clear. They provide a reference to check back on whether the objectives are met.
- External teams may require more detail on background context and on intended audiences and uses; internal teams may simply need to clarify the parameters of the assignment.

For managers of M/E activities
- They are a place to establish performance standards (e.g. reference to specific policies, standards).
- They are a means of building desired good practice into the process of the M/E activity (e.g. establishing a stakeholder consultation workshop in the methodology).
- They establish opportunities for quality control (e.g. presentation and review of intermediate products).

WHAT GOES INTO A PROGRAMME EVALUATION TOR?

The following can also be used for a project or activity-level evaluation.

Title
- Identify what is being evaluated. Use appropriate programme titles. Clarify the time period covered by the evaluation.

Background
- Briefly describe the history and current status of the programme, including objectives, logic of programme design or expected results chain, duration, budget, activities.
- Situate with reference to the organisation’s overarching country programme, as well as parallel or linked national programmes.
- Situate the important stakeholders, including donors, partners, implementing agencies/organisations.

Purpose of the evaluation
- Clarify why the programme is being evaluated.
- Describe how the evaluation process and/or results will be used and what value added they will bring.
- Identify the key users/target audiences.
- Situate the timing and focus of the evaluation in relation to any particular decision-making event (e.g. review meeting, consultation, planning activity, national conference) and/or the evolution of the programme.
Scope and focus

- An "objectives" format can be used with or instead of evaluation questions. Where both are used, one objective is usually discussed through a number of questions.
- List the major questions the evaluation should answer — they should relate to the purpose and be precisely stated so that they guide the evaluator in terms of information needs and data to collect. Group and prioritise the questions. They should be realistic and achievable.
- Specify evaluation criteria to be used given the evaluation's objectives and scope. Evaluations should use standard OECD/DAC criteria (relevance, efficiency, effectiveness, sustainability and impact) as well as additional criteria for evaluation of humanitarian response (coverage, co-ordination, coherence and protection). An explanation for the criteria selected and those considered not applicable should be given and discussed with the evaluation team.¹
- Evaluations of UNICEF-supported programmes should include two-additional criteria – the application of human rights-based approach and results based management strategies.
- Consider including a cost analysis of the programme. Good cost analysis strengthens results-based management and increases the utility of the evaluation.²
- Specify key policies and performance standards or benchmarks to be referenced in evaluating the programme, including international standards.

Existing information sources

- Identify relevant information sources that exist and are available, such as monitoring systems and/or previous evaluations. Provide an appraisal of quality and reliability.

Evaluation process and methods

- Describe overall flow of the evaluation process — sequence of key stages.
- Describe the overall evaluation approach and data collection methods proposed to answer the evaluation questions. An initial broad outline can be developed further with the evaluation team. Ultimately it should be appropriate and adequate providing a complete and fair analysis. The final TOR should define:
  - Information sources for new data collection
  - Sampling approaches for different methods, including area and population to be represented, procedures to be used and sampling size (where information is to be gathered from those who benefited from the programme, information should also be gathered from eligible persons not reached.)
  - The level of precision required
  - Data collection instruments
  - Types of data analysis
  - Expected measures put in place to ensure that the evaluation process is ethical and that participants in the evaluation – e.g. interviewees, sources — will be protected³
- Highlight any process results expected, e.g. networks strengthened, mechanisms for dialogue established, common analysis established among different groups of stakeholders.
- Specify any key intermediate tasks that evaluator(s) are responsible for carrying out, and a preliminary schedule for completion. Consider for example:
  - Meetings, consultation, workshops with different groups of stakeholders
  - Key points of interaction with a steering committee
  - Process for verification of findings with key stakeholders
  - Presentation of preliminary findings and recommendations.

Stakeholder participation

- Specify involvement of key stakeholders as appropriate providing a sound rationale — consider internal stakeholders, programme partners, donor representatives, etc. Roles might include liaison, technical advisory roles, observer roles, etc., or more active participation in planning and design, data collection and analysis, reporting and dissemination, follow-up.
- Specify expectations in terms of involvement of, or consultation with, primary stakeholders. Be clear about where they would participate, i.e. in planning and design, data collection and analysis, reporting and dissemination, and/or follow-up.

Accountabilities

- Specify the roles and responsibilities of the evaluation team leader and team members, as well as other stakeholders and advisory structures involved, e.g. steering committees. This section should clarify who is responsible for:
  - Liaison with the evaluation team
  - Providing technical guidance
  - Co-ordinating the stakeholders involved
  - Selection, orientation and training of team members, data collection assistants where applicable, interpreters
  - Approval of intermediate and final products
  - Capacity-building with stakeholders, national or other (a possible responsibility of the evaluation team).
- Specify the means to protect and limits to evaluators independence.
- Specify any concerns or restrictions related to conflicts of interest.

¹ For more on these criteria, see "Linking evaluation criteria with evaluation questions."
² For more on cost analysis, see Module 6, part 2.
³ For more on managing ethical considerations and protection of M&E participants, see "Ethical issues for field study – dealing with people" and "Children participating in research and M&E – Ethics and your responsibility as a manager"
Evaluation team composition

- Identify the composition and competencies of the evaluation team. This should follow from the evaluation focus, methods, and analyses required. Distinguish between desired and mandatory competencies, as well as whether competencies are required by the whole team or by certain members.
- Multidisciplinary teams are often appropriate. The qualifications and skill areas to be specified could include:
  - Areas of technical competence (sector, issue areas)
  - Language proficiency
  - In-country or regional work experience
  - Evaluation methods and data-collection skills
  - Analytical skills and frameworks, such as gender analysis
  - Process management skills, such as facilitation skills
  - Gender mix (not to be confused with gender analysis skills).

Procedures and logistics

- Specify as necessary logistical issues related to staffing and working conditions:
  - Availability and provision of services (local translators, interviewers, data processors, drivers)
  - Availability and provision of office space, cars, laptops, tape recorders, and procedures for arranging meetings, requirements for debriefings
  - Work schedule (hours, days, holidays) and special considerations such as in emergencies (e.g. often a 7-day work week is combined with R&R breaks)
  - Special procedures, for example on relations with press, security, evacuation in emergencies
  - Benefits and arrangements such as insurance (particularly in emergencies, consider hazard pay, war risk insurance)
  - Seasonal constraints, travel constraints/conditions and socio-cultural conditions that may influence data collection
  - Reporting requirements apart from products to be delivered (e.g. as accompanying invoices)

Products

- List products to be delivered, to whom and when. Consider:
  - The evaluation report
  - Completed data sets (filled out questionnaires or surveys)
  - Dissemination materials (newsletter articles, two-page summaries, presentation materials)
  - For UNICEF, evaluation consultants should be required to provide all of the information for the UNICEF CO update to the UNICEF Evaluation Database in the required format
  - Assessment of the evaluation methodology, including a discussion of the limitations.
- Specify the format for deliverables, including software, number of hard copies, translations needed and structure of the evaluation report. (See "UNICEF Evaluation Report Standards" and UNICEF Evaluation Technical Notes Series no. 3 "Writing a good Executive Summary".

Resource requirements

- Estimate the cost and prepare a detailed budget. Note the source of funds. Link the budget to the key activities or phases in the work plan. Cost estimates may cover items including:
  - Travel: international and in-country
  - Team member cost: salaries, per diem, and expenses
  - Payments for translators, interviewers, data processors, and secretarial services.
- Estimate separately any expectations in terms of time costs for:
  - Staff (before, during, after)
  - Other stakeholders, including primary stakeholders.
Writing a good Executive Summary

Primarily for key decision-makers who do not have time to read the full evaluation report, an Executive Summary should provide an overview of the essential parts of a report: a summary of the project/programme evaluated, the purpose of the evaluation, the methods used, the major findings and the recommendations. It should be very short — ideally two to three pages — and should “stand alone” (without requiring reference to the rest of the report).

The UNICEF’s Evaluation Database now lists the full Executive Summary for each report. The Executive Summary should be clear, simple and comprehensible to those not familiar with your programme, allowing database users to quickly grasp the important findings and recommendations.

By commitment to the Executive Board, Country Offices are requested to submit all completed evaluation reports to the Regional Office and to Headquarters [Evaluation Office] and all surveys and studies they believe are of significance in quality or findings. Also, Country Offices should use the format detailed below for all evaluations submitted to the Evaluation Office. It provides the needed information for effective use in the Evaluation Database and serves as a good stand-alone Executive Summary for inclusion in the report itself.

Depending on the audience for the report, it may be necessary to draft more than one Executive Summary or even different short dissemination pieces. In deciding whether to write a different Executive Summary than the standard format, consider your key audiences and the messages you want to bring to each one. The goal of the Executive Summary for the database is to share important findings and lessons with those outside of your programme in a concise manner.

The Executive Summary format below has two parts. The first part is standard bibliographical information, name of report, name of author, etc., and UNICEF-specific items, such as PIDB number, needed for report identification. The second part is the Executive Summary itself: background, purpose, methodology, findings, and recommendations. Both parts should be submitted to HQ. You may choose to use the second part as the Executive Summary inside of the report itself, but please still send both parts with the report to HQ to prevent confusion.
Format for Executive Summary

(Section 1 - Bibliographical Information)

Title: Usually includes the type of report and name of the project
Example: Evaluation of Early Childhood Development Programme

Author(s): Names of all the authors of the report in the following format: Last Name, First Initial; Last Name, First Initial
Example: Macom, X.; Pickett, W.

Institutions: Name of the institution contracted to implement the study, survey or evaluation. If a consultant is hired independently of an institution, this may be left blank
Example: London School of Hygiene and Tropical Medicine

Date: Year of publication

Region: Abbreviated name of UNICEF region. CEE/CIS, EAPRO, ESARO, MENA, ROSA, TACRO or WCARO

Country: Country name

Type: Survey, study or evaluation
♦ Survey: An assessment of the conditions of specified population group/s (children, women, adolescents) or public goods (health services, school, water system) at a point in time, e.g. MICS, KAP surveys; survey of quality of health services; Hygiene Practices survey

♦ Study: An investigation designed to improve knowledge about something (a problem or phenomenon) and understand its direct and underlying causes as well as its consequences on people or environment, e.g. A study of vulnerability of young people to illicit substance abuse

♦ Evaluation: A process to determine as systematically and objectively as possible the value or significance of a programme, project, policy or strategy, e.g. Evaluation of effectiveness and sustainability of UCI programme

Judgement on the value or significance is based on criteria such as relevance, effectiveness, efficiency, impact and sustainability. In rights-based programming, coverage (to what extent all social groups benefit) is another important evaluation criterion. Evaluations can use data from surveys and studies. Further, in evaluation of humanitarian assistance programmes, four additional criteria are recommended: coverage, coordination, coherence and protection.

Theme: Thematic area as defined in PIDB coding [add a button for a pull down list on website] The theme of an evaluation or study is the same as the theme of the related project or programme
**Partners:** Sponsoring organizations (funders and/or initiators)
UNICEF in collaboration with a donor, another international organization, a government ministry, a local research institution, etc.

**PIDB:** The actual PIDB code corresponding to the above theme is also necessary to establish a link between ProMS and the Evaluation Database in the future.

**Sequence Number:** Number assigned to the evaluation as reported in Annex A of the Annual Report.

**Follow Up:** One or two sentences about how the findings of the evaluation were used to improve programming.

**Languages:** Languages that the report is available in.

(Section 2 — Summary)

**Background:**
Brief information about the programme or project that the evaluation is related to, including the major stakeholders, partners and implementing organizations/agencies involved and their respective roles, as well as a brief description on how the evaluation came about.

**Purpose/Objective:**
A statement of why the assessment is needed, how it will benefit the program/project.
List your objectives, specifically, what you hoped to learn from the evaluation, questions to be answered by the evaluation.

**Methodology:**
A short description of the type of research methods used: sampling (describe the composition, location and number of people surveyed/interviewed/observed and how they were selected), data collection (survey, interviews, observation, desk review, etc.) Actual questions and surveys do not need to be included. Please include limitations if there are any significant drawbacks the audience should be aware of.

**Key Findings and Conclusions:**
Summary of significant findings and conclusions of the evaluation. All results do not need to be reported in full. Include important data and relevant, succinct conclusions drawn from findings. (If you have difficulty deciding what to include, a good rule of thumb is to look at those findings that led to your conclusions/recommendations; also key conclusions should be short answers to questions raised in the Objectives section.)

**Lessons Learned:**
(Optional, usually only given for thematic evaluations looking at a specific aspect beyond the level of one project or programme)
Recommendations that can be generalised beyond the specific case to apply to programs globally.

**Recommendations:**
Overall suggestions of how the project/program can be improved based on the findings.